DOCUMENT RESUME

ED 157 999 CE 016 798

AUTHOR Fox, Wayne L.; And Others
TITLE Individual Differences in Learning Performance: The Role of Aptitude.
INSTITUTION Human Resources Research Organization, Alexandria, Va.
PUB DATE [69]
NOTE 80p.; Not available in hard copy because of poor reproducibility

EDRS PRICE MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS Ability; Achievement Rating; Adult Vocational Education; *Aptitude; *Complexity Level; Enlisted Personnel; High Achievers; *Individual Differences; *Job Training; Learning Difficulties; *Low Ability Students; Low Achievers; Military Training; Performance Tests; Skill Development; *Task Performance; Teaching Methods; Vocational Aptitude
IDENTIFIERS Armed Forces Qualification Test

ABSTRACT

A study was designed to examine relationships of aptitude level to training performance on tasks of varying complexity and to determine if differential job training might be required. A complexity rated training battery with eight tasks (simple/choice monitoring, rifle assembly/disassembly, equipment preparation, map symbols, phonetic alphabet, and position plotting) was administered to 183 recruits selected through Armed Forces Qualification Test (AFQT) scores. Three homogeneous experimental groups were formed from subjects with different aptitude levels (HI-AFQT, MD-AFQT, and LO-AFQT). Results on simple and choice monitoring tasks revealed LO-AFQT subjects displayed poorer response times compared to HI-AFQT subjects. Rifle assembly/disassembly task differences were noted in favor of HI-AFQT subjects for mean response time and prompts. LO-AFQT subjects required more prompts, trials, and learning time than HI-AFQT subjects on the equipment preparation task. Wider within-group variability was noted for LO-AFQT subjects than other groups on map symbols and phonetic alphabet tasks. On the position plotting task LO-AFQT subjects required more training and displayed more within-group variability than HI-AFQT subjects. Findings suggest efficient training will depend on recognition of individual aptitude differences and instructional program designs compatible with individual learning rates and final performance capabilities. (CSS)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

- ED157999

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

MAR 1969

Fox

# INDIVIDUAL DIFFERENCES IN LEARNING PERFORMANCE: THE ROLE OF APTITUDE[1]

BEST COPY AVAILABLE

by

Wayne L. Fox, John E. Taylor and John S. Caylor

Human Resources Research Office, The George Washington University

## Introduction

Adult vocational education has received tremendous impetus in the past few years as an outgrowth of social and political pressures to provide adults with the training and education required for them to take a more productive and useful place in society. Not the least of this impetus has been provided by the Armed Services with the inauguration of Projects 100,000 and Transition.

In October 1966, the Department of Defense announced its decision to lower mental standards for induction and to eventually take into the military approximately 100,000 young men a year who were previously unacceptable. Indications are that these marginal aptitude trainees now constitute about 25 percent of the input to the U.S. Army training system.

Project Transition, the other Department of Defense program, is concerned with the retraining and utilization of men returning to civil life. In this project, men who are about to be discharged are being provided job training and counseling. Priority is being given to men with low school achievement and to those who have had no opportunity to acquire a civilian-related skill in service. Project Transition training and counseling are not restricted to men who enter military

CE 016 798

service under Project 100,000, but Transition is a primary source of benefits for these men.

The training job has become much more difficult with the advent of large numbers of marginal aptitude trainees into the Army training program, requiring much more effort on the part of instructors to bring these people with their typical histories of difficulty and frustration in school activities up to minimum acceptable standards.

The overall training problem, however, is not limited to the training of marginal aptitude personnel. In military, as in much civilian vocational instruction, it has been common practice to have students of all aptitude levels enter a given course together, use the same instructional materials, progress at the same rate, and leave the course together. While such civilian and military instruction tends to ignore individual differences in the classroom, civilian instruction seems directed toward the better students, typically allowing the poorer students to get on as best they can. The military instructor, however, in order to keep attrition rates at a minimum, orients his instruction toward the slower students, enforcing a slowed pace for the entire class which may well have an adverse effect upon the motivation and achievement of the higher ability students. Military training will inevitably be diluted in the effort to reach the increasing numbers of low aptitude students; consequently, resulting in a marked loss in motivation and achievement by the more apt as they become even more bored and restless than in the past (Montague & Showel, 1968).

With the recent announcement of the Department of Defense plan to take a higher proportion of college graduates into the military than in the past, an even greater range of student aptitudes, education and experience will be found in the typical military classroom.

It would seem almost axiomatic that a standard qualified training product cannot be achieved by putting widely differing students through a standard training mold. Only with the adaptation of training to individual differences can it be expected that students who differ extensively in their aptitude, education, and motivation will emerge with comparable levels of skill at the end of training.

There are a number of studies that indicate that individuals progress at different rates for different learning task (Cline, Beals & Seidman, 1956; Shay, 1961; Rigney, 1962); however, the interactions among aptitude level, training methods, and learning performance in practical training situations have been generally ignored.

## Research Problem

There is a definite need for research on the relationship of aptitude level to training performance on tasks of varying complexity to determine what, if any, differential training may be required for the efficient production of relatively standard occupationally-qualified individuals across a wide range of job-related tasks.

The research reported here was aimed at clarifying the relationship between aptitude level and learning performance on job-related training tasks in the interests of developing more effective training for individuals across the spectrum of aptitude.

## General Approach

The initial step in the research involved the development of a training task battery. The selection of tasks for inclusion in the battery was based on two criteria. The first was that the selected tasks should have elements in common with the skills and knowledges required in a large number of jobs. Examination of heavy density military occupations[2] yielded a number of tasks relevant to a variety of jobs.

The second criterion used for task selection was that tasks should be representative of several levels of complexity. Cagné's (1965) taxonomy of learning types served as a general guide for discriminating complexity differences among tasks. Gagné defined eight different types of learning which he ordered hierarchically from classical conditioning to problem solving. Our first criterion, that tasks be representative of practical jobs, prevented the selection of pure examples of each learning type as proposed by Gagné. The task battery, as finally selected, was composed of eight tasks which were roughly placed along a dimension of complexity as outlined in Table 1.

---

Insert Table 1 About here

---

Briefly, the task battery consisted of the following eight tasks discussed in order of complexity:

Simple Monitoring Task: In this task S was asked to perform a "watchkeeping" function which involved pressing a response lever when a stimulus light appeared on a display panel.

Choice Monitoring Task: This task was similar to the Simple Monitoring Task except that S had to respond to one of four possible stimuli by pressing one of four corresponding levers.

Rifle Assembly Task: Subjects were required to learn to assemble the M-14 rifle.

Rifle Disassembly Task: This task was similar to the Rifle Assembly Task above except that S was required to learn to disassemble the rifle.

Equipment Preparation Task: Subjects were to learn to perform 34 sequential steps necessary to prepare a missile for launching.

Map Symbols Task: Subjects were required to learn 26 map symbols.

Phonetic Alphabet Task: Subjects were required to learn the 26 letter international phonetic alphabet.

Position Plotting Task: Subjects were required to learn to plot the position of aircraft from information describing the range and bearing of the aircraft.

## Sample Attributes

Subjects were Army recruits with no previous military experience, which was defined to include ROTC, National Guard, and reserve duty. Before beginning formal processing at the U.S. Army Reception Station, Fort Ord, California, and typically within three days of having entered the Army, our Ss were taken to the HumRRO laboratory for one day of training on the experimental training tasks.

Subjects were selected solely on the basis of their scores on the Armed Forces Qualification Test (AFQT). The AFQT, a written mental aptitude test, is regarded as a general measure of trainability in military subjects. The AFQT is composed of 100 items divided into four equally weighted subtests: Vocabulary, Arithmetic Reasoning, Shop Mechanics, and Pattern Analysis. A score falling at the tenth centile on the AFQT standardization distribution is the statutory minimum set by Congress for acceptance into the military.

Three homogeneous and maximally different aptitude levels constituted the experimental groups. The high aptitude group (N=72) was defined by AFQT centile rank scores of 90-99; the middle group (N=50) by centile ranks 45-55; and the low aptitude group (N=81) by centile rank scores running from 10, the statutory minimum for Army service, through 21.

Table 2 presents the age distributions of the three aptitude groups.

---

Insert Table 2 about here

---

Education was recorded as the number of years of school completed, as reported by S and later reconciled with Army records. Table 3 presents the education distributions for the three aptitude groups.

---

Insert Table 3 about here

---

The relationship between AFQT score and amount of education is clear. Since the low aptitude group is slightly older than the high, the relationship of AFQT to education must be attributed not solely to age but to differential continuation in the school system. The high percentage of low aptitude men completing high school is surprising in that this group was not considered as sufficiently trainable for the Army until the recent relaxation of aptitude requirements under Project 100,000.

Table 4 presents the ethnic distributions for Ss at our three AFQT levels.

---

Insert Table 4 about here

---

Table 4 should be interpreted with caution. There are many factors producing seasonal and other systematic fluctuations in the socio-economic and ethnic characteristics of Army input. The time span and extent of our sampling is insufficient to warrant generalization of these ethnic

distributions. They are included here only as further description of our specific sample.

Reading level was assessed by use of the Gray Oral Reading Test, Form A, selected primarily for its wide range of reading levels. Subjects were required to read aloud short passages ranging from the simple level of "Look, Mother, look," through highly complex material. Distributions of scores, converted to grade levels by use of the published norms, are shown in Figure 1. In terms of reading level, the

---

Insert Figure 1 about here

---

three aptitude groups clearly differ. The low aptitude group was spread out almost evenly between grade level zero, complete inability to read, and the 11th grade level. None of the low AFQT sample reached the reading level attained by 71 percent of the middle AFQT group and by 94 percent of the high aptitude group.

Proficiency in elementary arithmetic was measured by a locally devised test consisting of simple items selected from the workbook used in a Supply Handlers Course. The 12 items of this test were:

1. 749 +213
2. 27862 +1865
3. 17625 - 739
4. 286542 -193663
5. 33 x33
6. 213 x78
7. 16/3200
8. 34/1156
9. 1/3+1/6+7/9=
10. 1/4-3/16=
11. 1/2x1/2=
12. 8/7÷16/14=

One point was awarded for each correct answer, giving a maximum possible score of 12. Table 5 presents the distributions of scores on

---

Insert Table 5 about here

---

each of the five parts of the arithmetic tests for the three aptitude level samples.

The Addition subtest was easy for these Ss and discriminated little between aptitude levels. On the other hand, comparison of the percentage of Ss failing all items on a subtest shows values, for the high and low aptitude groups respectively, of 2% vs. 43% for the Subtraction subtest, 4% vs. 21% for the Multiplication subtest, 0% vs. 55% for the Division subtest, and 0% vs. 51% for the Fractions subtest. Comparison of men getting all items in a subtest correct shows an equally extreme and sharp contrast between aptitude levels.

No standardization data are available to permit expression of these scores in terms of school grade level or in comparison with the general population in this age group. Nevertheless these arithmetic items do not appear to be of great difficulty. As with reading, the relationship between arithmetic ability and AFQT is clear.

The cumulative percentage distributions of the three aptitude groups on the total score of the arithmetic test are shown in Figure 2. In this

---

Insert Figure 2 about here

---

figure a point on the curve shows, on the ordinate at the left, the percentage of men in that aptitude group scoring at or below the score indicated on the baseline. Thus, 81% of the low aptitude group scored 6 or below out of a possible 12 points, while only 14% of the middle AFQT group and none of the high aptitude groups scored no higher than 6.

Similarly, 59% of the high group, 36% of the middle group, and only 5% of the low aptitude group achieved total scores of 11 or 12 (these being the differences between the percentage of Ss scoring 10 or lower and 100%).

Extensive psychometric data were extracted from records of routine Army testing conducted for purposes of classification and assignment.

These data were summarized in terms of the percentage of men in each AFQT group scoring above 100 in the Army standard norm score distribution. Unlike the centile rank norm scores, running from 1-100, which are used with the AFQT, the scores of the Army classification tests reported in Table 6 have been transformed to a normal distribution with

---

Insert Table 6 about here

---

with an arithmetic mean of 100 and standard deviation of 20 for the mobilization population.

Table 6 presents the percentage of individuals in each AFQT group scoring above 100 on each of the subtests of the Army's classification and assignment battery. Data are based on 79 Ss in the low AFQT groups, 29 in the middle AFQT group, and 71 cases in the high AFQT group.

The frequency distributions on which each row of data in Table 6 are based show some slight variability from one subtest to another. As a set, however, they can be fairly summarized as typically and consistently showing a wide range of scores for the middle AFQT groups with scores spanning a substantial portion of the full range from the bottom score of the low AFQT group to the top of the high group. With equal consistency, there is trivial or no overlap between the classification test score ranges for the high and low AFQT groups.

In summary, the pattern of psychometric test performance for the three AFQT levels is of a consistency rarely encountered. The high AFQT group performs outstandingly well in all areas, the low AFQT performs correspondingly consistently poorly on the average, and the middle AFQT group falls in an intermediate position, with great individual variability. For many of the abilities measured by our instruments and the Army Classification Battery, this is not surprising for the same attributes are being measured by both tests and our findings serve only to validate and refine the AFQT measures. Thus, the poor performance of the low AFQT groups on the experimental measures of reading and arithmetic merely corroborates the Verbal and Arithmetic subtests of the AFQT and makes distressingly clear the low level of ability in these fundamental skills.

## General Methodology

As stated above, this study was undertaken to learn the effects of wide differences in aptitude on the learning of a range of laboratory tasks representative of job training. Since a major focus was on those of men of marginal aptitude, instructional methods were selected to maximize the low aptitude Ss opportunity to learn. The selection of instructional methods and learning conditions was established, on a judgmental basis, without regard for considerations of cost and effort, or for efficiency for the middle and high aptitude groups.

Where practical, instruction was automated to insure standardization and clarity, using audio-visual presentation including slides and video tape. Verbal instructions were given in simple language with ample pictorial examples. All training was conducted individually with an instructor present to give prompts, answer questions, and to provide immediate knowledge of results after each response. Material was

presented in the smallest possible integral segments. Instructions were repeated or reviewed, as appropriate, and practice was provided on each trial. In short, training procedures were tailored, within our judgment, to give the lower aptitude student the best possible opportunity to learn.

Within the seven hours of training time available for any one S, it was not possible to attempt all eight training tasks. Moreover, since training was individually administered and continued for a variable time until a criterion level of performance was reached, no fixed set of tasks could be scheduled. Accordingly, different individuals performed different sets of training tasks, accounting for fluctuations in sample size among tasks.

The substantial differences in sample size for different categories of data were a function of the design and conduct of the study. The choice of tasks to be administered to each of the three new Ss available each day for training was a function of several factors. The training tasks themselves were implemented at different times during the five month period of data collection. Since all training was conducted individually, and run for whatever time was required to reach criterion, assignment to training tasks was necessarily contingent upon the availability of subject time, training equipment, and trainer personnel.

Initially, training tasks were assigned so as to insure an approximately equal N of both the high and low aptitude groups for each task. Midway through the data collection, training of middle aptitude Ss was initiated, at which time assessment of reading and arithmetic proficiency was instituted for all three aptitude groups. Training of the middle aptitude group was concentrated in the five more difficult and complex tasks, judged more appropriate to their aptitude level, because of

training time limitations. Information on age, education, AFQT, and ethnic category was obtained on all individuals regardless of which set of training tasks they undertook.

For the high and low aptitude groups, assignment of training tasks fell in between the extremes of having each S learn five tasks or having each S learn only a single task. Most high and low aptitude Ss were trained in only a few of the eight experimental tasks. For the extreme aptitude groups, this assignment procedure approached the condition of independent random sampling for each task.

## Simple and Choice Monitoring Tasks

The Simple and Choice Monitoring Tasks are considered to be representative of a number of tasks requiring visual surveillance or watchkeeping activity. Included would be tasks performed by personnel whose main function is to detect and react to a signal (e.g., switchboard operators, radar operators, and control panel monitors). These tasks represent the simplest level of complexity included in the task battery.

### Method

Seventeen HI-AFQT Ss and 15 LO-AFQT Ss performed the Simple Task, and 19 HI-AFQT and 19 LO-AFQT Ss performed the Choice Task. Different Ss performed in each task.

The apparatus consisted of a stimulus panel, appropriate response levers, and assorted recording devices located in a sound-deadened cubicle. The same basic apparatus was used in both tasks varying only in the number of available response levers.

Figure 3 shows the stimulus panel which was divided into four sections

---

Insert Figure 3 about here

---

labelled A, B, C, and D. Each section contained two rows of three white lights separated by a single red light. A single lever, as shown in Figure 3 for the Simple Task, and four levers corresponding to the ABCD panel sections for the Choice Task, were situated in front of the stimulus panel.

The 24 white lights accompanied by a loud clicking noise were programmed to come on one at a time in random sequence at the rate of five per second.

The four red lights were programmed to come on one at a time at intervals ranging from 15 to 205 seconds in 10 second steps. The resulting 20 interstimulus intervals were randomized with the restriction that each interval appear only once during each task. Each red stimulus light came on a total of five times in random order.

The onset of a red stimulus light automatically shut down the program controlling the white lights so that when a red light was on the white lights were off. The clicking noises that accompanied the white lights, however, continued at the previous rate of five per second during the time that the red stimulus light was on. The red light remained on until the appropriate lever was pressed. Pressing the lever also activated the program controlling the white lights and signalled the beginning of the next interstimulus interval.

Upon entering the testing room, S was told that he was to monitor an "Operations Control Panel." It was explained that the flashing white lights indicated that different pieces of information were entering the central control room located next door, that when too much information was processed through the panel the white lights went out, and one of the red lights came on indicating that no further information could be processed until the "overloaded" circuits were reset. It was the S's task to reset the circuits by pressing the single lever in the Simple

Task, or one of the four levers in the Choice Task. The importance of speed in making his response was emphasized. The stimulus panel was then activated and S made a practice response. If there were no questions, the panel was reactivated and the first of the 20 trials began.

An instructor was present in the cubicle at all times to monitor a timer which automatically recorded the response time measured from the onset of the red stimulus light to the activation of the lever. Each presentation of a red stimulus light constituted a trial.

## Results

Skewed response time distributions on several trials due to a few very long (i.e., 10 seconds) response times dictated the use of medians rather than means as the more representative measure of central tendency.

Median response times are presented in Figure 4 for both high and low

---

Insert Figure 4 about here

---

aptitude Ss on the Simple and Choice Tasks. The 20 trials on each task are grouped into blocks of four trials each. As expected, longer response times were recorded in the Choice Task than in the Simple Task.

Median scores for each S were computed across the 20 trials and then summed across Ss in each aptitude group to provide the data presented in Table 7. The high aptitude Ss made significantly faster responses than

---

Insert Table 7 about here

---

the low aptitude subjects on both tasks.

These results indicate that at relatively simple levels of task complexity, as exemplified by the Simple and Choice Monitoring Tasks,

low aptitude Ss displayed poorer performance in terms of response time when compared to higher aptitude Ss.

## Rifle Assembly and Disassembly

Every man entering military service, except for those conscientious objectors assigned to the medical services, is trained to assemble and disassemble the M-14 rifle. These tasks are fixed-procedure motor tasks which have elements in common with a variety of tasks performed in many jobs. In addition to assembly/disassembly procedures, fixed-procedure motor tasks are involved in the setting up and operation of the whole spectrum of mechanical and electronic equipment. Such fixed procedural tasks require a series of motor responses which must be performed in a specified order. These rifle tasks were judged to represent a higher level of task complexity than the simple monitoring tasks discussed above.

### Method

Seventy-six Ss, 23 HI-AFQT, 30 MD-AFQT, and 23 LO-AFQT, were trained to assemble the M-14 rifle, and 30 Ss, 18 HI-AFQT and 20 LO-AFQT, were trained to disassemble the rifle. Different Ss performed each task.

Figures 5 and 6 show the classroom where training was conducted. The

---

Insert Figures 5 & 6 about here

---

room included a closed-circuit TV screen, table, standard Army Disassembly Mat for the M-14 rifle, an M-14 rifle, and an instructor.

Subjects were led step-by-step through either correct assembly or disassembly procedure by a qualified rifle instructor appearing on video tape. Both demonstrations (Assembly and Disassembly) were divided into seven distinct steps corresponding to the separate parts of the rifle. At the completion of each step by the demonstrator, the video tape was

stopped and S was directed by the instructor beside him to perform the same assembly or disassembly step on the rifle provided him. This procedure was followed for each of the seven steps until the rifle was completely assembled or disassembled. Completion of a cycle of the seven steps constituted a training trial. After each training trial, S undertook a test trial during which he attempted to assemble or disassemble the rifle completely without the aid of the video tape demonstration. Training and test trials were continued until S had received a minimum of three training trials, and he stated that he needed no further training trials. Test trials were continued until S showed no further improvement on three consecutive trials in time-to-assemble or -disassemble score.

During both training and test trials a prompt was given by the instructor after 30 seconds when S showed no progress, or at any time when aid was requested. A maximum of three instructor-initiated prompts was given at 30 second intervals for each assembly or disassembly step, the last prompt given after approximately 90 seconds had elapsed and always ending in instructor-guided assembly or disassembly of the part, or conclusion of the step.

Subjects were told to work as rapidly as possible but not to skip steps or attempt short-cut procedures. The instructor recorded (a) total time-to-assemble or -disassemble the rifle (part times were accumulated during training trials), and (b) the number of prompts given during training and test trials. Following inspection of the data, a response time criterion was determined for each task. Final trial scores for all Ss were used to construct a frequency distribution and the tenth centile was selected as the criterion performance level. This method of criterion selection assured that only extreme cases would be unable to surpass the

criterion. The criterion scores obtained in this way were 70 seconds for the assembly task and 50 seconds for the disassembly task.

## Results

The results for the Rifle Assembly and Disassembly tasks are presented separately below.

### Rifle Assembly Task

Rifle assembly mean response times are presented in Figure 7 for the

---

Insert Figure 7 about here

---

high, middle, and low aptitude groups. Data are presented only through Trial 11 as the N per aptitude group began to decrease differentially after this point as individuals reached asymptotic performance and completed the task. An analysis of variance indicated that the overall AFQT effect was significant ($F[2,73] = 23.27$, $p = <.001$), as was the trials effect ($F[10,730] = 477.56$, $p = <.001$).

The large initial difference among groups became systematically smaller as practice continued, resulting in a significant AFQT by trials interaction effect ($F[20,730] = 21.98$, $p = <.001$). Apparently there was a significant differential ability to profit from the first instruction trial. This was true even though none of the Ss had previous experience with the M-14 rifle. The reduction in the magnitude of the differences among groups as a function of practice, however, was not great enough to eliminate the AFQT effect by Trial 11. Comparisons of mean response times on the last trial (Trial 11) indicated significant differences among all three groups.

The group means presented above do not permit examination of individual differences in criterion attainment within aptitude groups.

To facilitiate such a comparison the trial by trial cumulative percentage of Ss meeting the 70 second performance criterion for the rifle assembly task is presented in Figure 8. Inspection of the curves at any trial will

---

Insert Figure 8 about here

---

show the percent of Ss in each group who met criterion at the end of that trial. Note, for instance, that by Trial 7, 73% of the Ss in the HI-AFQT group, 50% of the Ss in the MD-AFQT group, and 26% of the Ss in LO-AFQT group had bettered the 70 second criterion. Every man in the HI-AFQT group had bettered the criterion by the ninth trial, while after the 15th trial, a substantial percentage of men in the other two groups had yet to attain criterion.

The mean number of prompts received by each aptitude group is presented in Table 8. The majority of prompts for all three groups

---

Insert Table 8 about here

---

occurred during the first three trials and no further prompts were required after the conclusion of the last training trial (Trial 5).

Rifle Disassembly Task

The rifle disassembly task yielded data similar to that presented for the rifle assembly task. Rifle disassembly mean response times are presented in Figure 9 for the high and low aptitude groups. Data are

---

Insert Figure 9 about here

---

presented only through Trial 10 as the number group began to decrease differentially as individual Ss reached asymptotic performance and completed the disassembly task.

Analysis of variance indicated that the overall AFQT effect was significant ($F[1,36] = 26.23$, $p = <.001$), as was the trials effect ($F[9,324] = 159.21$, $p = <.001$), and the AFQT by trials interaction effect ($F[9,324] = 10.85$, $p = <.001$). As in the rifle assembly task, the systematic reduction in the difference between groups (which led to the significant AFQT by trials interaction) was not great enough to eliminate the AFQT effect. Here, too, apparently there was differential ability to profit from the initial instruction. Comparison of mean response times on the last trial (Trial 10) again yielded a significant difference between high and low aptitude groups ($t[36df] = 3.01$, $p = <.01$).

Curves for the cumulative percentage of Ss reaching the response time criterion on each trial are presented in Figure 10. Note that on the

---

Insert Figure 10 about here

---

sixth trial 72% of the HI-AFQT group and 15% of the LO-AFQT group had bettered the 50 second performance criterion. All but two Ss in the high aptitude group had bettered the criterion by Trial 8, while one-half of the low aptitude group had still to attain criterion at that point.

The mean number of prompts received by the high and low aptitude groups during the disassembly task appear in Table 9. As in the rifle

---

Insert Table 9 about here

---

assembly task, the majority of prompts occurred on the first three trials with no further prompts required after the conclusion of the last training trial (Trial 5).

In summary, analysis of the rifle assembly and disassembly data indicated that there were significant differences in the acquisition of skill in these two fixed-procedure motor tasks among Ss of different aptitude levels. Clearly, all Ss learned to perform their assigned task (assembly or disassembly of the rifle), but there were differences among aptitude groups in the time taken to assemble or disassemble the rifle (mean response time), the amount of individualized instruction needed to meet task requirements (prompts), and the number of trials required to attain a specified criterion.

All differences were in favor of the higher aptitude Ss. Low aptitude Ss took about twice as long as those of high aptitude to acquire minimum proficiency on the tasks, and required more than twice as much individualized help. On the rifle assembly task, where data were available on middle aptitude Ss, their scores were seen to fall midway between the high and low aptitude groups on all the measures.

Large differences were found among Ss in their ability to profit from the initial instruction. This was true even though all Ss were supposedly starting out even, as none had reported any previous experience with weapons of this kind. On both tasks, these large initial differences in response time were gradually reduced as a function of practice; however, final response time performance levels were shown to favor the higher aptitude Ss.

Cumulative percentage curves showing the percent of Ss in each group meeting the performance time criteria on each trial showed a wide variability in individual performance within aptitude groups. There were

some low aptitude Ss who were able to attain criterion with a minimum of training and practice, while others did not reach criterion on the last reported trial. On the other hand, there were a few high aptitude Ss who did relatively poorly on the tasks.

## Equipment Preparation

The immediately preceding tasks, Rifle Assembly and Disassembly, were representative of fixed-procedure motor tasks. The Equipment Preparation Task is also a fixed-procedure task but emphasizes learning a series of verbal responses. Verbal procedure tasks are found in many jobs that require the use of checklists, either explicit or implicit, in the setting up and operation of equipment (e.g., engine trouble shooting, setting fuses and preparing charges, preparing aircraft for flight, checking out radios). This task was considered to be at roughly the same level of complexity as the preceding assembly/disassembly tasks.

### Method

Twenty HI-AFQT Ss, 25 MD-AFQT Ss, and 21 LO-AFQT Ss participated in the Equipment Preparation Task.

The task consisted of a 34-step procedure intended to prepare a missile for launch. The procedure was performed on a specially designed training device which simulated a guidance system control panel.

Each of the 34 steps in the procedure required a single response which consisted of touching an electronic stylus to one of approximately 100 electrical contacts located on the screen of the training device. The training device, S, and the instructor were all together in a small soundproof room. The correct procedure was demonstrated by the instructor who carefully explained the purpose of each of the 34 steps and pointed out the corresponding electrical contacts on the training device. At the conclusion of the demonstration, S was given a written checklist

containing a short description of each of the 34 steps in proper sequence. He was given the opportunity to read through the checklist (or have it read to him) and ask questions before the beginning of the first training trial. The checklist was available throughout the task.

In addition to providing the checklist during each training trial, the instructor gave verbal and spatial prompts when S failed to make a correct response. Seven seconds was allowed for each correct response. A correct response was followed by the cycling of the training device to present the next step in the sequence. Nothing happened when S made a wrong response. There was no limit to the number of responses that could be made during the seven second interval. Failure to make a correct response within the time limit resulted in an automatic spatial prompt by the training device itself. An arrow appeared on the screen pointing to the correct electrical contact. In addition to the spatial prompt by the machine, the instructor provided an oral prompt from a standardized list corresponding to each of the 34 steps.

The training device automatically recorded the number of prompts and the total time to complete the 34 steps. Each completion of the 34 step sequence constituted a training trial. Training was continued until S required fewer than five prompts per trial or until he had received a total of 15 trials.

## Results

The mean number of prompts per trial for the three aptitude groups is presented in Figure 11. Inasmuch as some Ss in each group reached

---

Insert Figure 11 about here

---

criterion and completed the task in fewer trials than other members of the group, the N per trial was held constant by continuing to record criterion scores for those Ss completing the task before the slowest group member.

There were clearly differences in the mean number of prompts per trial required by the three aptitude groups. The HI-AFQT group mean bettered the criterion on the third trial and the MD-AFQT group bettered the criterion on the fifth trial. The LO-AFQT group mean, however, did not surpass the criterion until the final trial (Trial 15). The mean number of prompts required to attain criterion performance or to complete Trial 15 was 22.9 for the HI-AFQT group, 42.9 for the MD-AFQT group, and 133.0 for the LO-AFQT group. These differences were highly significant ($F[2,64] = 40.81$, $p=<.001$).

Table 10 shows the mean number of trials and the mean amount of time

---

Insert Table 10 about here

---

in minutes required by each group to better the criterion of four or fewer prompts per trial. Five LO-AFQT Ss failed to meet criterion and were arbitrarily assigned scores based on 15 trials. Thus, the means and standard deviations are underestimated for the LO-AFQT group. The low aptitude group required at least three times as many trials and four times as much time to attain criterion (or complete 15 trials) as did the high aptitude group, and at least twice as many trials and more than twice as much time as the middle aptitude group.

A comparison between high and middle aptitude groups showed that the HI-AFQT group required significantly fewer trials ($t = 3.78$, $p = <.001$)

and significantly less time ($t = 3.79$, $p=<.001$) to attain criterion than did the MD-AFQT group.

Figure 12 presents the cumulative percentages of Ss in each aptitude

---

Insert Figure 12 about here

---

group reaching criterion on each trial. Note that by Trial 5, 95% of the HI-AFQT group and 69% of the MD-AFQT group had bettered the criterion, while by the same trial only 10% of the LO-AFQT group had achieved criterion performance. After 15 trials, 24% of the LO-AFQT group had not yet attained criterion proficiency. There was also a substantial percentage of low aptitude Ss who reached criterion in the same number of trials as some of the slower high aptitude Ss (29% by Trial 7).

In summary, low aptitude Ss required six times as many prompts, three times as many trials, and at least four times as long as high aptitude Ss to reach criterion on the task. The middle aptitude Ss required approximately one-third as many prompts, one-half as many trials, and less than half as much time as the low aptitude Ss to attain criterion. Variability among trainee groups was inversely related to aptitude level with low aptitude Ss showing the greatest variability in performance.

## Map Symbols and Phonetic Alphabet

The learning of map symbols and the international phonetic alphabet are examples of tasks involving multiple discrimination. The Map Symbols Task required learning to associate words with symbols, while the Phonetic Alphabet Task required learning to associate letters of the alphabet with corresponding phonetic equivalents. Examples of other tasks of this nature include learning hand and arm signals, cooking times and temperatures for food, part names and equipment nomenclature, and color coding.

These tasks were considered to occupy an intermediate level of complexity in the task battery.

## Method

Eighty Ss, 25 HI-AFQT, 30 MD-AFQT, and 25 LO-AFQT, participated in the Map Symbols Task; and 74 Ss, 21 HI-AFQT, 29 MD-AFQT, and 24 LO-AFQT, participated in the Phonetic Alphabet Task.

### Map Symbols

This task consisted of learning 26 commonly used map symbols. Each symbol appeared on a 5" x 8" study card with its appropriate name, and an artist's representation of the thing, place, or event which was represented by the symbol. Figure 13 is an example of the study card used for the "artillery" symbol.

---

Insert Figure 13 about here

---

After a pretest to determine the S's familiarity with the symbols, he was shown the 26 study cards, one at a time, by an instructor who identified each symbol and read the symbol name aloud. The S was then given the study cards and instructed to learn them as he would be tested on them again shortly. Three minutes were allowed to study the 26 symbols.

At the conclusion of the three-minute study period, S was given an answer sheet containing the symbols in scrambled order. He was instructed to identify each symbol. Oral responses were accepted from those Ss who were unable to spell or write legibly. All other Ss responded by writing the appropriate symbol name in a blank space adjacent to each symbol. The instructor recorded correct responses as they were written (or given orally) on a separate record sheet. At the conclusion of the testing period, the instructor indicated the errors that were made and told S the

correct responses. Trials were continued until two successive errorless trials were recorded (the S had correctly named the 26 symbols on two successive trials) or until he had completed 12 trials.

Phonetic Alphabet

The Phonetic Alphabet Task consisted of learning the 26 international phonetic equivalents to the alphabet. All of the 26 letters and their corresponding phonetic equivalents were printed on a single 8½" x 11" card in correct alphabetical sequence: A-ALFA, B-BRAVO, C-CHARLIE, etc.

The general procedure for the Phonetic Alphabet Task followed that as outlined above for the Map Symbols Task. After a pretest, the instructor read each letter and its appropriate phonetic equivalent aloud. Subject was then given the study card and allowed to study for three minutes. After the study period, he was presented with an answer sheet containing the letters of the alphabet in correct alphabetical sequence. Subject was to respond with the correct phonetic equivalent. As in the Map Symbols Task, oral responses were accepted from those Ss unable to spell or write clearly. Errors were corrected by the instructor immediately following the test period. Trials were continued until S correctly responded with all 26 phonetic equivalents on two successive trials, or until he had completed eight trials.

## Results

The results for the Map Symbols and Phonetic Alphabet Tasks are presented separately below. The second criterion trial served only as a test trial to assure that learning had in fact occurred; it was not included as a separate trial in the data analysis.

Map Symbols

The mean number of correct responses per trial for the Map Symbols Task are presented in Figure 14 for the three aptitude groups. Inasmuch as some

---

Insert Figure 14 about here

---

Ss in each group attained criterion and completed the task in fewer trials than other members of the group, the N per trial was held constant by continuing to enter scores of 26 for those Ss who had reached criterion before the slowest group member.

As shown in Figure 14, every S in the HI-AFQT group had attained criterion performance by the third trial. The slowest Ss in the MD-AFQT group required seven training trials to reach criterion, while not all LO-AFQT Ss had completed the task by Trial 12.

A comparison of Map Symbols pretest means indicated that the HI-AFQT group (pretest mean = 1.7) did significantly better on the pretest ($p = <.05$) than did the MD-AFQT (pretest mean = .8) and the LO-AFQT (pretest mean = .5) groups. These pretest differences, though statistically significant, were hardly large enough to account for the subsequently observed differences in acquisition rate among groups noted in Figure 14.

Means and standard deviations of trials-to-criterion scores are presented in Table 11. Five Ss in the LO-AFQT group failed to meet

---

Insert Table 11 about here

---

criterion and were assigned scores based on 12 trials. The mean and standard deviation presented in Table 11 are thus underestimates of the performance of this group. It can be seen, however, that the LO-AFQT group took about twice as many trials to attain criterion as the MD-AFQT group, and that the MD-AFQT group took about twice as many trials to reach criterion as the HI-AFQT group.

The cumulative percentage of Ss reaching criterion per trial presented in Figure 15 permits comparisons within groups as well as among groups.

---

Insert Figure 15 about here

---

Note, for instance, that on Trial 3 all of the HI-AFQT Ss had reached criterion, while 63% of the MD-AFQT Ss and only 12% of the LO-AFQT Ss had learned all of the symbols by Trial 3. Note also that the low aptitude group exhibited much greater variability than the other two groups in attaining the criterion. Almost one-half of the low aptitude Ss had learned the symbols by Trial 5, while at the other extreme, 21% of them had not learned the symbols by Trial 12 when training was stopped. The variability within the other two groups is much less pronounced.

Phonetic Alphabet

The mean number of correct responses per trial for the Phonetic Alphabet Task is presented in Figure 16 for the three aptitude groups. As in the

---

Insert Figure 16 about here

---

Map Symbols Task data analysis, the N per group was held constant by entering scores of 26 for those Ss in each group completing the task before the slowest group member. These curves (Figure 16), when compared to those of the Map Symbols Task (Figure 14), indicate that learning rates for the two tasks were quite similar and that the Phonetic Alphabet Task was apparently relatively less difficult for the middle and low aptitude groups. Two important task differences possibly having an effect were: (1) the additional cues provided by the matching of the alphabet letter and the first letter of the phonetic equivalent, and (2) the constant serial position

of stimulus-response pairs throughout the Phonetic Alphabet Task. The HI-AFQT groups seemed not to be as susceptible to these task differences as they appeared to perform both tasks with about equal facility.

A comparison of pretest means on the Phonetic Alphabet Task indicated no significant differences among groups.

Table 12 contains the means and standard deviations of trials-to-criterion

---

Insert Table 12 about here

---

scores for the Phonetic Alphabet Task. Analysis of variance indicated significant differences among groups. Between-group comparisons indicated that there were no significant differences between the high and middle aptitude groups, and further, that both of these groups learned in a significantly fewer number of trials than the low aptitude groups.

The cumulative percentage of Ss reaching criterion per trial is presented in Figure 17. The cumulative percentage curves are much alike

---

Insert Figure 17 about here

---

for the high and middle aptitude groups, both differing markedly from that of the low aptitude group. As in the Map Symbols Task, all of the HI-AFQT Ss had attained criterion by the third trial. Almost all of the MD-AFQT Ss had also reached criterion by Trial 3, while fewer than half of the LO-AFQT Ss had done so. Although not as pronounced as in the Map Symbols Task, wider variability is again evidenced within the LO-AFQT group.

## Position Plotting

The Position Plotting Task involved the learning and application of principles. Here, S had to learn the concepts of range and bearing and

apply them in an intersection problem to plot the position of a target. Similar tasks are common to a variety of target acquisition or plotting jobs which make use of maps and grid coordinates. The Position Plotting Task represents the highest level of complexity which was included in the task battery.

## Method

Twenty-four HI-AFQT, 23 MD-AFQT, and 24 LO-AFQT Ss participated in the Position Plotting Task.

The task involved learning to plot the position of aircraft from data describing the target's bearing and range relative to Ss position. Instruction in plotting techniques was given using a coordinated audio-tape/35 mm. slide program presented via closed circuit TV. Specifically, the instructional period consisted of first, the definition of bearing, including examples, and practice on using the concept to determine direction from a given point. Second, the definition of range was given and S practiced determining distance from a given point. Finally, S practiced using both concepts to make plots of the position of aircraft. A plotting board like that shown in Figure 18 was used for the practice problems. An instructor

---

Insert Figure 18 about here

---

was present to provide help, direct practice, and answer questions throughout the instructional sequence.

At the conclusion of training, Ss were required to make 10 plots on the plotting board. They were given the bearing, in degrees, and the range, in miles, of an aircraft, and were required to draw an "X" on the board at the point of intersection. During the testing periods, Ss were allowed seven seconds to make each plot. After each plot, the instructor provided

immediate knowledge of results by indicating on the plotting board the correct point of intersection using the plotting techniques previously shown. After each block of 10 plots, the instructional sequence, except for introductory material, was repeated. Training was continued until S was able to make nine out of 10 correct plots on two successive blocks (trials) or until he had undergone a total of 10 trials.

## Results

The mean correct plots per trial are presented in Figure 19 for the

---

Insert Figure 19 about here

---

three aptitude groups. The second criterion trial was used only to assure that S could perform the task reliably and was not used in the data analysis. The N per trial for each group was held constant by supplying criterion scores for those Ss reaching criterion before the slowest member of each group.

The means for the middle and high aptitude groups were above criterion on the very first test trial following instruction, although not all Ss in these two groups had attained criterion performance. The low aptitude group, however, needed much additional training to approach criterion performance.

The means and standard deviation of the trials-to-criterion scores are presented in Table 13. Analysis of variance indicated the overall

---

Insert Table 13 about here

---

differences were highly significant; however, there was no significant difference between the high and middle aptitude groups. The differences

between these two groups and the low aptitude group are quite large, even though six of the 24 LO-AFQT Ss failed to meet criterion and their scores are thereby underestimated.

The cumulative percentage of Ss meeting criterion provides a picture of individual performance. It can be seen in Figure 20 that three-fourths

---

Insert Figure 20 about here

---

of the Ss in the high aptitude group could perform the task after the initial presentation of the plotting instructions, while fewer than 10 percent of the low aptitude Ss were capable of criterion performance following the initial instructional sequence. After 10 trials of instruction and practice, there was still one-fourth of the LO-AFQT Ss who had not attained criterion performance. All but one of the middle aptitude Ss were able to make accurate plots by the third training trial, with the remaining individual requiring seven trials.

The high and middle aptitude groups had little trouble mastering the Position Plotting Task. The low aptitude group required much more training time than the higher aptitude groups. Greater variability was also displayed by the low aptitude group in that a few Ss were able to master the plotting task early in training, while a substantial number failed to attain criterion performance within the 10 trials allotted for instruction and practice.

## Summary, Discussion and Implications of Findings

The objective of the work reported here was to determine the extent to which mental aptitude (as measured by the AFQT) is related to the acquisition of skills and knowledge such as required in a variety of occupational specialties.

Current technology has little information for use in designing instructional programs to accommodate the entire spectrum of aptitude, and practically nothing is known specifically about the engineering of instruction for those of lower ability. If educators hope to develop programs that will be effective for all individuals, a necessary first step is that of assessing the relationship of student aptitude to learning performance.

Our sample of 183 Army recruits was selected so as to constitute three homogeneous non-overlapping aptitude groups. These three groups were found to be consistently and highly different on a variety of measures, as summarized below.

Entry Level Characteristics

As elaborated above, whether we examined psychometric data (Army Classification Battery) or scholastic achievement (educational level completed, reading, arithmetic proficiency), we found the high aptitude Ss to be decidedly superior to the low aptitude Ss, with the middle aptitude groups scoring in an intermediate range. It was somewhat surprising to discover that so many of the low aptitude Ss had completed high school (61%). To the extent that high school graduation is based on academic achievement no higher than that indicated by AFQT, reading, and arithmetic scores of the low aptitude sample, our data suggest that educational attainment as a variable for predicting performance will become of decreasing usefulness.

Laboratory Findings

Data were collected on a battery of training tasks selected to be representative of the skills and knowledge found in a variety of jobs. The eight training tasks that composed the task battery varied in complexity from relatively simple stimulus-response association and procedural learning tasks, to the more complex learning of multiple-discriminations, and

concepts and principles. Instructional methods and learning conditions were established, on a judgmental basis, to afford the lower aptitude individual the best possible opportunity to learn without regard for consideration of cost or effort, or for efficiency for the middle and high aptitude groups.

The data indicated that large and consistent differences, related to aptitude, were found on some seven measures employed to assess learning performance in the laboratory learning tasks. These measures included response time, trials and time-to-criterion, prompts, number of correct responses per trial, and estimates of variability. Low aptitude Ss displayed poorer performance than higher aptitude Ss on all these measures.

The finding that the aptitude groups differed in their response time scores in the monitoring tasks is not taken as having great practical significance in itself. However, if such relatively simple tasks show these differences, it may well be that similar tasks requiring more complicated response patterns would find the aptitude groups separated by greater time intervals. The major significance of these findings is taken to be that high aptitude individuals are apparently able to make these kinds of perceptual motor responses faster than the lower groups. In field situations (monitoring, target detection, aircraft identification) where tasks are more complex, compounded by confusion, or overlaid by other tasks, such speed of response differences may well have important practical impact.

On the more complex training tasks, differences in training time required were of such magnitude (factors of 4 and 5 when comparing lows with highs) that they are interpreted as having considerable significance. The practical significance of these findings may loom largest in the training of individuals of low aptitude to perform such tasks as equipment preparation (a long procedural task) and position plotting (a complex task

involving concepts and principles) under more realistic training conditions. The equipment task employed here was only one-third as long as the actual missile checkout procedure from which it was adapted. Further, typical procedural training of this nature does not routinely provide the student with written checklists to follow. The plotting task we employed required S to learn only range and bearing, just two of the many variables involved in position plotting and target acquisition.

It is thus highly conceivable that under operational, rather than laboratory training conditions, the development of proficiency in performing the more complex tasks included in the task battery might prove to require more time in training for lows over highs than the factors of 4 and 5 found here. One wonders if current lockstep instructional practice ever does succeed in bringing low level individuals to high levels of proficiency in such important tasks.

Those tasks which provided measures of the amount of prompting and instructional guidance required to attain criterion performance yielded results which are seen to be of considerable import. Here, much like the learning time measures, comparison of lows with highs showed difference factors of 5 and 6. This finding is seen to have implications for the amount of guidance and instructor intervention needed by students of differing aptitude levels. The following observations go beyond the data to be sure, but low aptitude students probably require frequent access to an instructor, they probably require considerable guidance and coaching, and they probably are penalized severely as class size becomes large. This is not to say whether the "instructor" is a live human, a self-teaching program, or a computer.

Although we selected groups highly homogeneous on AFQT, the performance of these groups on our learning tasks was not equally homogeneous. The high

aptitude group showed least variability and all of them reached criterion [illegible]mance level on every task. The middle aptitude group displayed sligh[illegible] more variability and on only two tasks did a few of them fail to reach criterion. Variability of learning performance was notably greater in the low aptitude group. On almost every task there were a few low aptitude Ss whose performance matched that of the middle and high groups; similarly on almost every task there were some low aptitude Ss who failed to reach criterion, and the remainder were strung over a wide range of performance. The individual identified as of "low aptitude" is not necessarily a slow learner on all tasks. Those individuals who are of marginal aptitude on the AFQT are not qualitatively alike; they constitute a heterogeneous group. Apart from low general aptitude, a low AFQT score can arise from many factors unrelated to later learning performance; language difficulty; unfamiliarity with testing procedures; and motivation, mood, and physical condition at the time of testing. An important problem to be solved lies in the differential sorting of low aptitude personnel. There is a need for devising ways to identify the faster learners, and their areas of promise, among those labelled as being of low aptitude. The design of this study, which was directed to the comparison of the performance of extreme aptitude groups on a large number of different learning tasks, precluded any analysis of individual consistency of behavior.

One remaining aspect of these findings deserves discussion at this point, that being the interaction of aptitude effects with training trials. In most instances, the data showed large differences among aptitude groups on the early trials, with this difference decreasing systematically over trials. In some tasks all groups eventually reached the same level of performance; in other tasks they appeared to reach different asymptotes. Apparently, there was a differential ability to profit from initial

instruction. High, and in many cases middle aptitude Ss, were able to master the material in just a few training trials. One can speculate that the higher aptitude groups were just brighter and could therefore assimilate and process information faster, that they had learned how to learn and thus had a set conducive to rapid learning, or, that they were able to adapt better to novel situations (recall that the laboratory battery consisted of a number of short, unrelated, discrete tasks, no one of which consumed more than an hour and a half). As an individual progressed through the task battery, he was faced by a number of novel situations. It could be hypothesized that rather than reflecting differences in aptitude or basic ability to learn, these observed differences were in large part reflecting nothing more than simple adaptability to novel situations. This hardly seems tenable, however, in view of the consistency of the laboratory findings with performance on the various other measures.

The general pattern of these findings, across a variety of measures sampling performance upon entry into the Army and on the task battery, is one of striking consistency. Such large and consistent aptitude effects have not been reported in the technical literature heretofore, probably because the typical study of aptitude effects has varied aptitude over a limited range only (college sophomores form the typical subject population). The present study, in order to assess performance over the entire meaningful range of aptitude, selected samples of Ss so that rather pure groups representing high, middle, and low ranges of aptitude could be studied.

The consistency of these findings is all the more striking considering the number and variety of Ss that were involved in the study. Rather than being based upon the performance of fixed and possibly unique groups of Ss on whom many repeated measures were taken, these data are based upon the

performance of many different individuals who underwent many different combinations of tasks.

Our data could not be analyzed to determine whether task complexity itself relates to performance. Though the eight laboratory tasks were selected so as to vary in complexity, there was no attempt made to scale them along a dimension. Thus, though there is general agreement that Map Symbols (multiple discrimination) is more complex than Rifle Disassembly (motor procedure) we cannot estimate how much more complex it actually is; we cannot determine their relative positions on a quantitative scale of complexity.

Further, though the tasks were ordered as to complexity, task difficulty was not controlled. Thus, though Equipment Preparation was less complex than Map Symbols, it was much more difficult to learn because of its length.

It is reasonable to hypothesize that lower aptitude individuals would have less and less learning success as task complexity increases. These data do not allow us to test such a hypothesis. We can say only that, regardless of task complexity, mental aptitude is strongly related to performance.

Implications of Findings

The well aired issues of homogeneous grouping and the multi-track curriculum represent classical approaches to solving the problem of lock-step instruction by providing for individual differences.

What may be the largest and fastest moving area of research in public education today is that which is concerned with the individualization of instruction. Programmed Instruction, Computer Assisted Instruction, Individually Prescribed Instruction--all approaches to the individualization of instruction--are predicated on the existence of differential abilities to profit from instruction.

Under Project 100,000 a wide variety of training is being designed, re-engineered, or in some fashion modified to accommodate a wider range of student aptitude. Unfortunately, training technology does not specify how to go about this; it is not known which training strategies are most appropriate for which aptitude levels. However, early results being reported from some four or five Army courses being modified specifically for low aptitude students indicate that these individuals are helped considerably when the training is geared to their level to insure their assimilation of the material.

Considering the previous research, the experience to date of Project 100,000 and the pattern of findings reported from the present study, there can be no doubt that student aptitude is a potent variable that must be recognized in the conduct of instruction. But student aptitude differences account for only part of the picture. The findings of studies of instructional method, the results reported from redesigned Project 100,000 courses, and the expanding literature on the individualization of instruction, indicate that instructional method is likewise a potent factor. It is clear that no one single instructional method is effective across all aptitude levels; that individuals of differing aptitude levels require instructional methods that match their aptitude and, of course, motivation and experience. The college graduate and the low aptitude student simply cannot be reached by the same instructional vehicle. However, these same studies confirm the meagerness of our current knowledge on how to go about designing appropriate training strategies for the various levels of aptitude. This is a serious gap in the technology of training and education.

## Conclusions

The laboratory findings discussed in this report, in conjunction with the entry level findings, lead us to the following conclusions:

1. Mental aptitude, as measured by performance on the Armed Forces Qualification Test, relates consistently to a variety of important psychometric and operational criteria, which include:

a. Performance on the Army's psychometric tests for classification and assignment.

b. Scholastic achievement as indicated by scores on reading and arithmetic tests, and by school grade level completed.

2. Learning performance is directly related to aptitude level. This relationship holds across a variety of training tasks which vary in complexity. This relationship is demonstrated by an array of response measures which show that:

a. In some tasks aptitude groups differ in rate of learning only.

b. In some tasks aptitude groups differ both in rate of learning and in final level of performance.

c. In simple response tasks aptitude groups differ in speed of response.

d. The time required to train low aptitude Ss and high aptitude Ss to comparable levels differs substantially.

e. The learning performance of middle aptitude groups is more similar to that of high aptitude groups than it is to low aptitude groups.

f. Performance variability relates inversely to aptitude level. Not all individuals labelled as being of low aptitude are slow learners on all tasks; on each task, a few show performance typical of the middle and high aptitude groups.

g. The requirement for instructor guidance and prompting is related inversely to aptitude level.

The relationship of aptitude to the aforementioned measures is a consistent and powerful one with important implications for the efficient

conduct of training. High and middle aptitude groups generally outperform low aptitude groups by a wide margin. These findings, considered in the light of related studies, imply that the efficient training of men at all levels of aptitude will depend upon (a) the recognition of individual differences in aptitude, and (b) the design of instructional programs which are compatible with individual differences in learning rate and final performance capability.

References

Cline, V. B., Beals, A. R., & Seidman, D. Evaluation of four and eight weeks basic training for men of various intelligence levels. Paper presented at the American Psychological Association meetings, Chicago, Sept. 1956.

Gagne, R. M. The Conditions of Learning. New York: Holt, Rinehart, & Winston, 1965.

Montague, E. K., & Showel, M. A review of the general supply course (MOS 76A10) at Fort Ord, California. Staff paper, Human Resources Research Office, Division No. 3, July 1967.

Rigney, J. W. Potential uses of computers as teaching machines. In J. E. Coulson (Ed.), Programmed learning and computer-based instruction. New York: Wiley, 1962, pp. 155-170.

Shay, C. B. Relationship of intelligence to step size on a teaching machine program. Journal of Educational Psychology, 52, 1961, 98-103.

## Footnotes

1. The research reported in this paper was performed at HumRRO Division No. 3 (Recruit Training), Presidio of Monterey, California, under Department of the Army contract with The George Washington University. The contents of this paper do not necessarily reflect official opinion or policies of the Department of the Army.

2. Supply Clerk, Wheel Vehicle Mechanic, Cook, Light Vehicle Driver, Personnel Clerk, Wireman, etc.

Table 1

Ordering of Tasks Along a Dimension of Complexity

| | Description of Learning Requirements | Task |
|---|---|---|
| Complex | Requires learning of concepts and principles and their application in a problem situation. | Position Plotting |
| | Requires multiple discrimination of words and symbols (serial or paired-associate learning). | Map Symbols<br>Phonetic Alphabet |
| | Requires learning of fixed procedures, either verbal or motor (chaining of verbal or motor responses). | Equipment Preparation (verbal procedure)<br>Rifle Assembly<br>Rifle Disassembly (motor procedure) |
| Simple | Requires association of stimulus and response. | Choice Monitoring<br>Simple Monitoring |

Table 7

Age Distributions of Aptitude Groups

| | Age | | | | | |
|---|---|---|---|---|---|---|
| AFQT | 18-19 | 20 | 21 | 22 | 23-25 | N |
| 90-99 | 11% | 36% | 15% | 18% | 20% | 72 |
| 45-55 | 10% | 63% | 7% | 7% | 13% | 29 |
| 10-21 | 18% | 21% | 20% | 8% | 33% | 81 |

Table 3

Education Distribution of Aptitude Groups

| | Years of Schooling Completed | | | | | |
|---|---|---|---|---|---|---|
| AFQT | 10 or Fewer | 11 | 12 | 13 | 14 or More | N |
| 90-99 | 0% | 1% | 12% | 24% | 63% | 72 |
| 45-55 | 13% | 0% | 42% | 22% | 23% | 29 |
| 10-21 | 23% | 16% | 53% | 5% | 3% | 81 |

Table 4

Ethnic Distribution of AFQT Levels

| | Ethnic Designation | | | | |
|---|---|---|---|---|---|
| AFQT | Anglo | Mexican | Negro | Oriental | N |
| 90-99 | 90% | 0% | 2% | 8% | 72 |
| 45-55 | 79% | 11% | 6% | 4% | [illegible] |
| 10-21 | 38% | 27% | 20% | 15% | 81 |

Table 5

Distribution of Part Scores on Math Test

(Percent Achieving Each Sub-Score)

| | Add | | | Subtract | | | Multiply | | | Divide | | | Fractions | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 0 | 1 | 2 | 0 | 1 | 2 | 0 | 1 | 2 | 0 | 1 | 2 | 3 | 4 | |
| AFQT | | | | | | | | | | | | | | | | | | N |
| 90-99 | 0 | 4 | 96 | 2 | 11 | 87 | 4 | 34 | 62 | 0 | 9 | 91 | 0 | 4 | 17 | 17 | 62 | 46 |
| 45-55 | 0 | 3 | 97 | 7 | 37 | 56 | 3 | 30 | 67 | 3 | 23 | 74 | 13 | 20 | 17 | 13 | 37 | 30 |
| 10-21 | 0 | 23 | 77 | 43 | 32 | 25 | 21 | 47 | 32 | 55 | 23 | 22 | 51 | 28 | 17 | 2 | 2 | 47 |

Table 6

Percentage of AFQT Groups Scoring Above 100 on Army Classification Tests

| | AFQT | | |
|---|---|---|---|
| Subtest | 10-21 | 45-55 | 90-99 |
| Verbal | 4 | 76 | 100 |
| Arithmetic | 1 | 75 | 100 |
| Shop Mechanics | 37 | 59 | 96 |
| Pattern Analysis | 9 | 66 | 100 |
| Army Clerical Speed | 32 | 72 | 89 |
| Automotive Information | 33 | 59 | 93 |
| Mechanical Aptitude | 14 | 72 | 97 |
| Electrical Information | 19 | 45 | 93 |
| General Information | 13 | 69 | 96 |
| Classification Inventory | 28 | 45 | 73 |
| Army Radio Code Aptitude | 24 | 72 | 94 |

Table 7

Overall Group Means ($\bar{X}$) and Variances ($s$)[a]

| | Simple Task[b] | | Choice Task[c] | |
|---|---|---|---|---|
| | $\bar{X}$ | $s$ | $\bar{X}$ | $s$ |
| HI-AFQT | .50 | .04 | .68 | .10 |
| LO-AFQT | .55 | .07 | .79 | .12 |

[a]Comparable analyses were performed using mean response time per trial and yielded essentially the same results reported for the median response times.

[b]$t = 2.63$, $p = < .02$

[c]$t = 3.03$, $p = < .01$

Table 8

Rifle Assembly Means ($\bar{X}$) and Standard Deviations ($s$) of Prompts Per Group

| Group | $N$ | $\bar{X}$ [a] | $s$ |
|---|---|---|---|
| HI-AFQT | 23 | 6.4 | 4.1 |
| MD-AFQT | 30 | 11.3 | 6.0 |
| LO-AFQT | 23 | 16.6 | 7.9 |

[a] $F(2,73) = 15.72$, $p = < .001$

Table 9

Rifle Disassembly Means ($\bar{X}$) and Standard Deviations ($s$) of Prompts Per Group

| Group | $N$ | $\bar{X}$ [a] | $s$ |
|---|---|---|---|
| HI-AFQT | 18 | 5.4 | 3.7 |
| LO-AFQT | 20 | 12.2 | 5.4 |

[a] $F(1,36) = 19.86$; $p = < .001$

Table 10

Means ($\bar{X}$) and Standard Deviations ($s$) of Trials-to-Criterion and Time (in minutes)-to-Criterion Scores

| Aptitude Group | N | Trials-to-Criterion[a] $\bar{X}$ | Trials-to-Criterion[a] $s$ | Time-to-Criterion[b] $\bar{X}$ | Time-to-Criterion[b] $s$ |
|---|---|---|---|---|---|
| HI-AFQT | 20 | 3.3 | 1.2 | 11.3 | 5.6 |
| MD-AFQT | 26 | 5.0 | 1.7 | 18.4 | 6.8 |
| LO-AFQT | 21 | 10.4[c] | 3.7[c] | 45.5[c] | 18.5[c] |

[a] $F(2,64) = 49.05$, $p = < .001$

[b] $F(2,64) = 51.00$, $p = < .001$

[c] Subjects who did not attain the criterion of fewer than five prompts were arbitrarily assigned scores based on 15 trials. All high and middle Ss met the criterion. Five LO-AFQT Ss failed to meet the criterion. Thus, the mean and standard deviation for the LO-AFQT group are underestimated.

Table 11

Map Symbols Means ($\bar{X}$) and Standard Deviations ($s$) of Trials-to-Criterion Scores

| Aptitude Group | N | $\bar{X}$ [a] | $s$ |
|---|---|---|---|
| HI-AFQT | 25 | 1.7 | .8 |
| MD-AFQT | 30 | 3.3 | 1.6 |
| LO-AFQT | 25 | 6.2[b] | 2.4[b] |

[a] $F(2,77) = 42.35$, $p = < .001$

[b] Subjects who did not attain the criterion of errorless performance were arbitrarily assigned scores based on 12 trials. Five LO-AFQT Ss failed to meet criterion, thus the means and standard deviations of the LO-AFQT groups are underestimated.

Table 12

Phonetic Alphabet Means ($\bar{X}$) and Standard Deviations ($s$) of Trials-to-Criterion Scores

| Aptitude Group | $N$ | $\bar{X}$ [a] | $s$ |
|---|---|---|---|
| HI-AFQT | 21 | 1.9 | .8 |
| MD-AFQT | 29 | 2.1 | .9 |
| LO-AFQT | 24 | 4.0[b] | 1.9[b] |

[a] $F(2,71) = 20.75$, $p = < .001$

[b] Subjects who did not attain errorless performance were arbitrarily assigned scores based on eight trials. Only one S failed to attain criterion. As he was in the LO-AFQT group, the mean and standard deviations for this group are slightly underestimated.

Table 13

Means ($\bar{X}$) and Standard Deviations ($s$)

of the Number of Trials-to-Criterion

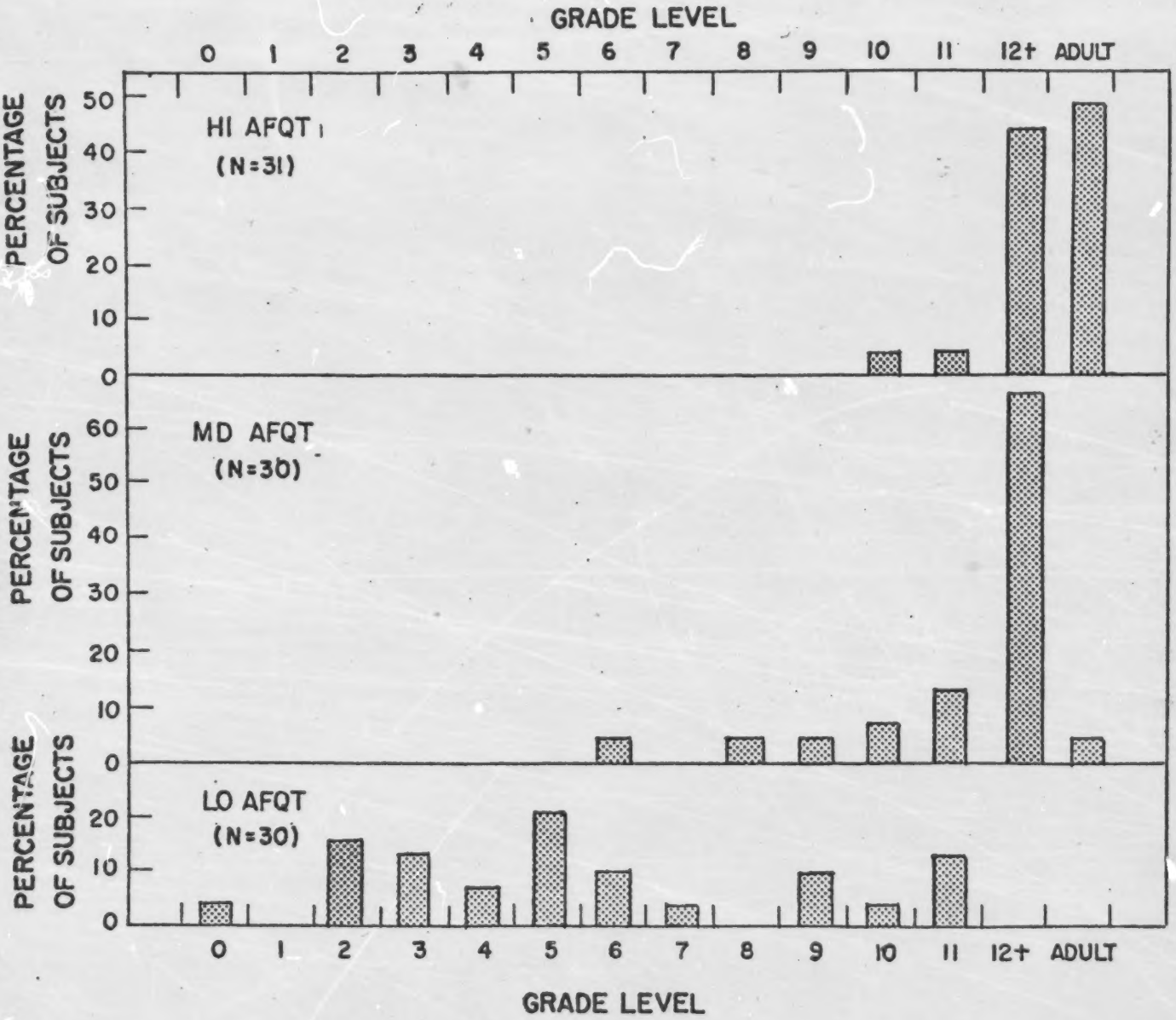
| Aptitude Group | $N$ | $\bar{X}$ [a] | $s$ |
|---|---|---|---|
| HI-AFQT | 24 | 1.3 | .6 |
| MD-AFQT | 28 | 1.7 | 1.3 |
| LO-AFQT | 24 | 5.2[b] | 3.3[b] |

[a] $F(2,73) = 25.83$, $p = < .001$

[b] Subjects who did not meet the criterion of nine out of 10 correct plots were arbitrarily assigned scores based on 10 trials. Six of the LO-AFQT Ss failed to attain criterion by the conclusion of Trial 10, thus the mean and standard deviation reported for the LO-AFQT are underestimated.

Figure Captions

1. Gray Oral Reading Test: Percentage Distribution by Grade Level
2. Arithmetic Test: Cumulative Percentage of Subjects at or Below Indicated Total Score
3. Monitoring Tasks Stimulus Panel
4. Simple and Choice Monitoring Tasks: Median Response Times by Aptitude Level
5. Rifle Assembly/Disassembly Task Apparatus
6. Rifle Instructor and Trainee
7. Rifle Assembly: Mean Response Time Per Trial
8. Rifle Assembly: Cumulative Percentage of Subjects Reaching Criterion Per Trial
9. Rifle Disassembly Mean Response Per Trial
10. Rifle Disassembly Cumulative Percentage of Subjects Reaching Criterion Per Trial
11. Equipment Task: Mean Number of Prompts Per Trial
12. Equipment Task: Cumulative Percentage of Subjects Reaching Criterion Per Trial
13. Artillery Map Symbols Study Card
14. Map Symbols: Mean Number Correct Per Trial
15. Map Symbols: Cumulative Percentage of Subjects Reaching Criterion Per Trial
16. Phonetic Equivalents: Number Correct Per Trial
17. Phonetic Equivalents: Cumulative Percentage of Subjects Reaching Criterion Per Trial
18. Position Plotting Board
19. Position Plotting: Mean Number of Correct Plots Per Trial.
20. Position Plotting: Cumulative Percentage of Subjects Reaching Criterion Per Trial.


GRAY ORAL READING TEST: PERCENTAGE DISTRIBUTION BY GRADE LEVEL
GRADE LEVEL
0 1 2 3 4 5 6 7 8 9 10 11 12+ ADULT
PERCENTAGE OF SUBJECTS
HI AFQT
(N=31)
50 40 30 20 10 0
MD AFQT
(N=30)
60 50 40 30 20 10 0
LO AFQT
(N=30)
20 10 0
0 1 2 3 4 5 6 7 8 9 10 11 12+ ADULT
GRADE LEVEL

59 FIGURE 1.

ARITHMETIC TEST: CUMULATIVE PERCENTAGE OF SUBJECTS AT OR BELOW INDICATED TOTAL SCORE

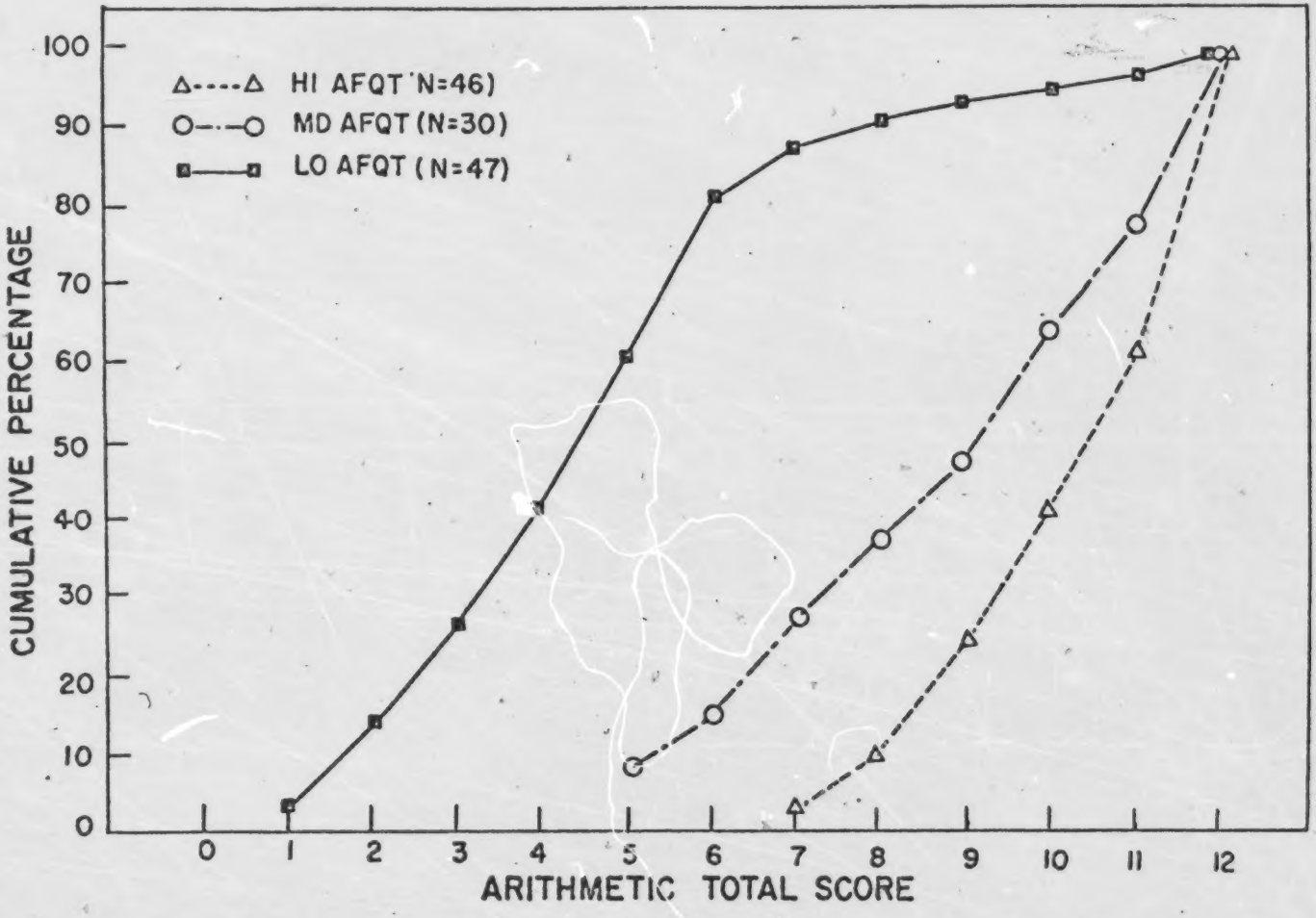
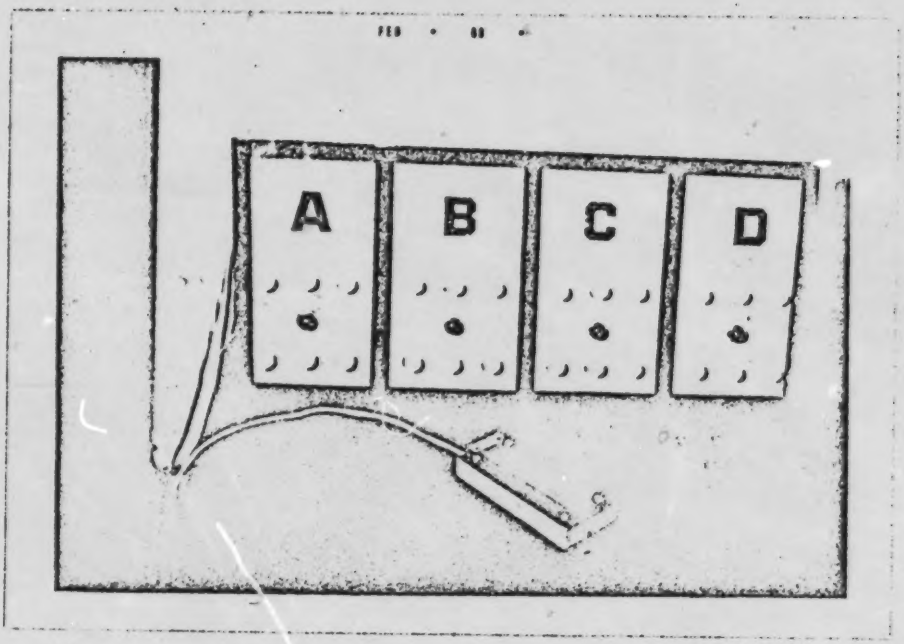



FIGURE 2






MONITORING TASKS STIMULUS PANEL

SIMPLE AND CHOICE MONITORING TASKS: MEDIAN RESPONSE TIMES BY APTITUDE LEVEL

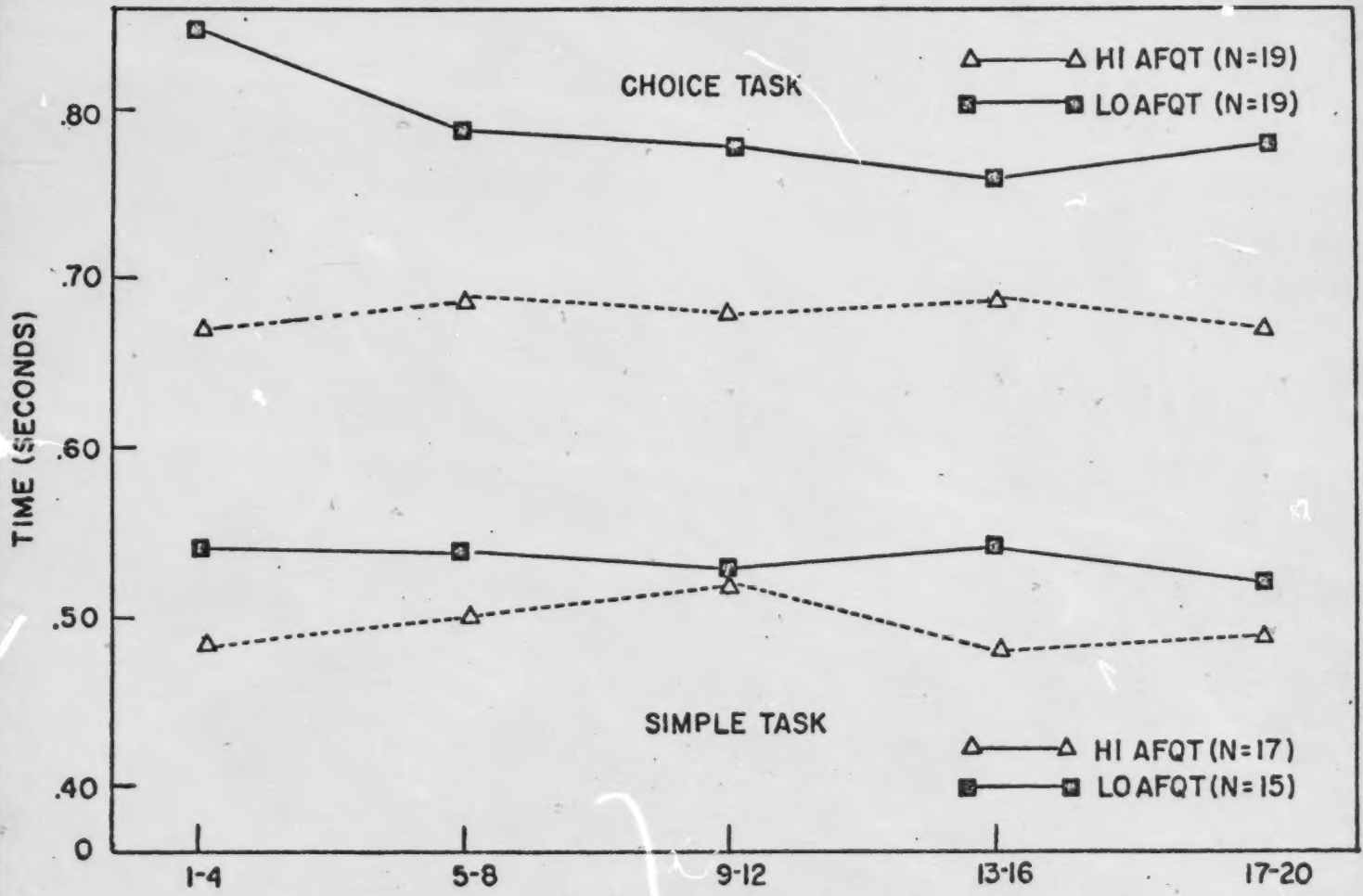
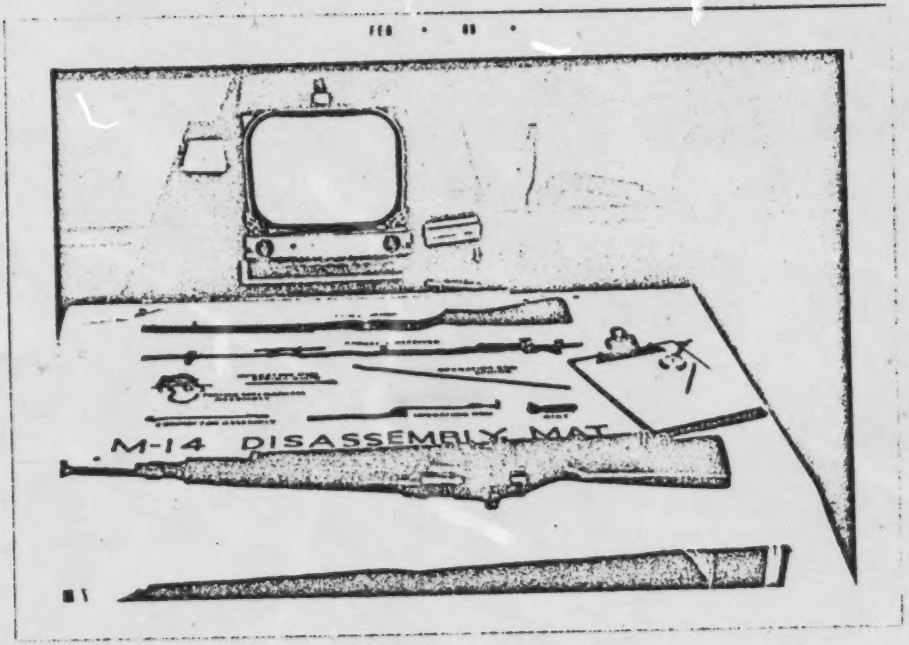



FIGURE 4






RIFLE ASSEMBLY/DISASSEMBLY TASK APPARATUS

FIGURE 5

RIFLE INSTRUCTOR AND TRAINEE

FIGURE 6

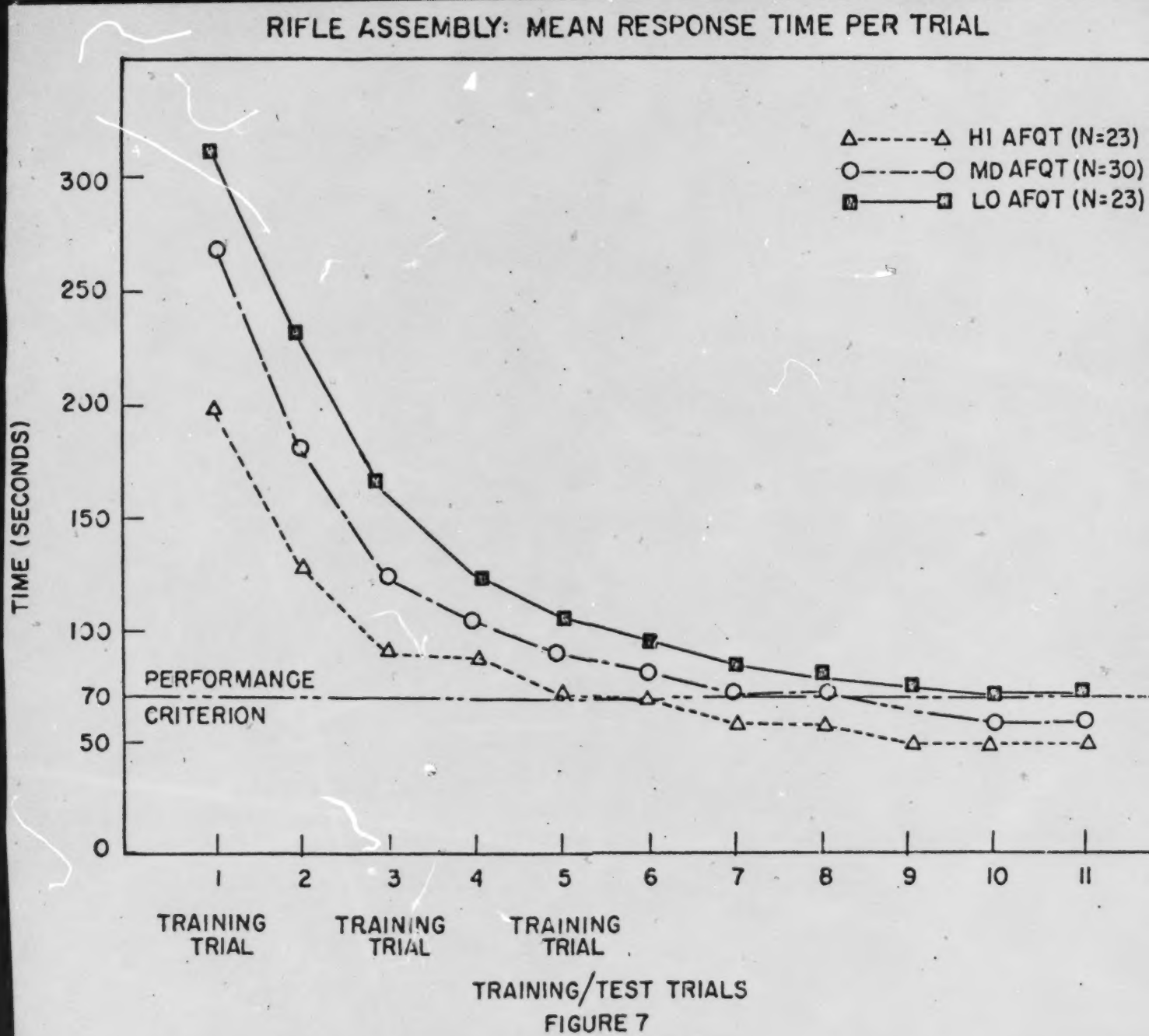

RIFLE ASSEMBLY: MEAN RESPONSE TIME PER TRIAL
HI AFQT (N=23)
MD AFQT (N=30)
LO AFQT (N=23)
TIME (SECONDS)
300
250
200
150
100
70
50
0
PERFORMANCE
CRITERION
1
2
3
4
5
6
7
8
9
10
11
TRAINING TRIAL
TRAINING TRIAL
TRAINING TRIAL
TRAINING/TEST TRIALS

FIGURE 7

RIFLE ASSEMBLY: CUMULATIVE PERCENTAGE OF SUBJECTS REACHING CRITERION PER TRIAL

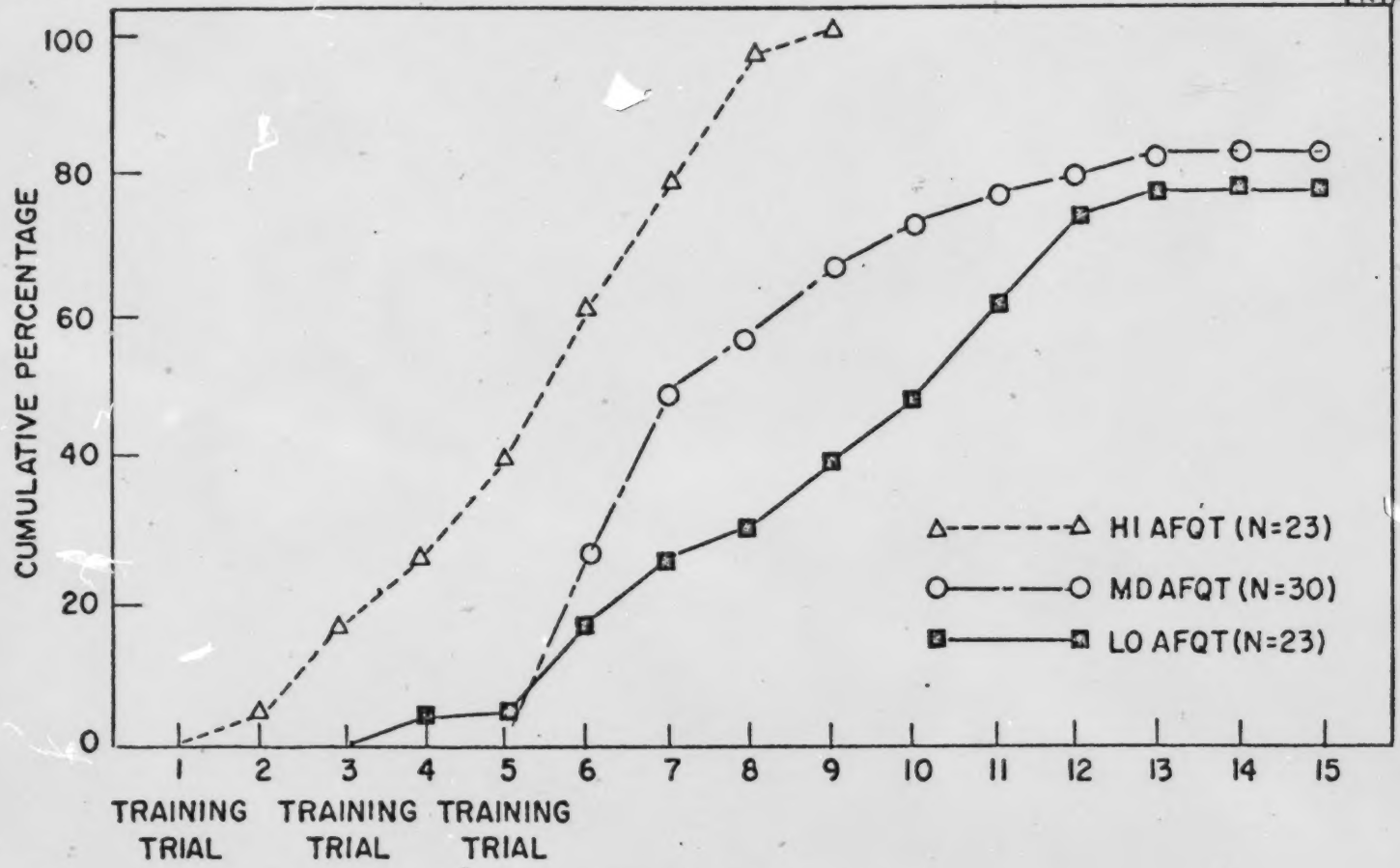



FIGURE 8

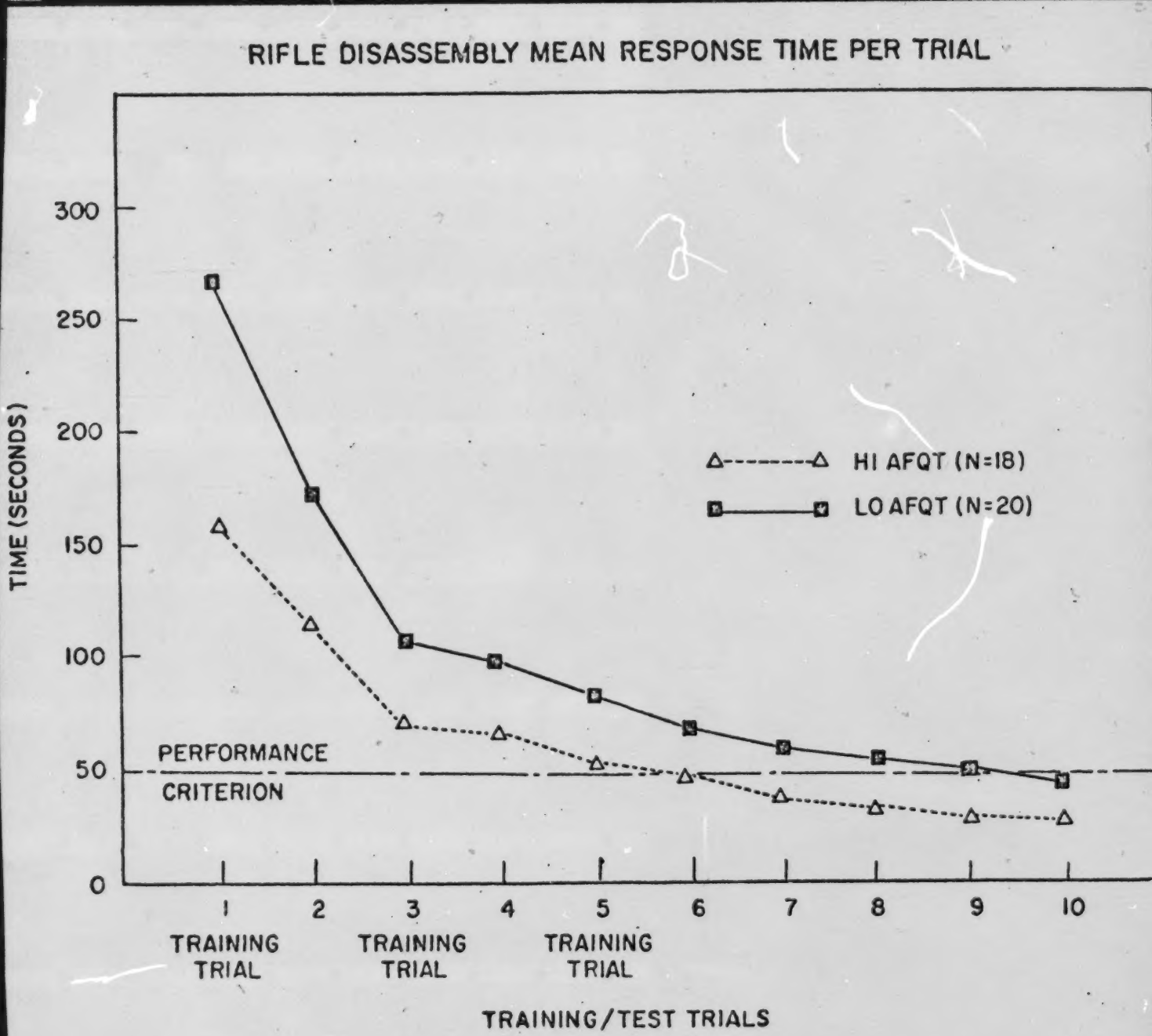

RIFLE DISASSEMBLY MEAN RESPONSE TIME PER TRIAL
TIME (SECONDS)
300
250
200
150
100
50
0
HI AFQT (N=18)
LO AFQT (N=20)
PERFORMANCE
CRITERION
1
2
3
4
5
6
7
8
9
10
TRAINING TRIAL
TRAINING TRIAL
TRAINING TRIAL
TRAINING/TEST TRIALS

 FIGURE 3

RIFLE DISASSEMBLY CUMULATIVE PERCENTAGE OF SUBJECTS REACHING CRITERION PER TRIA

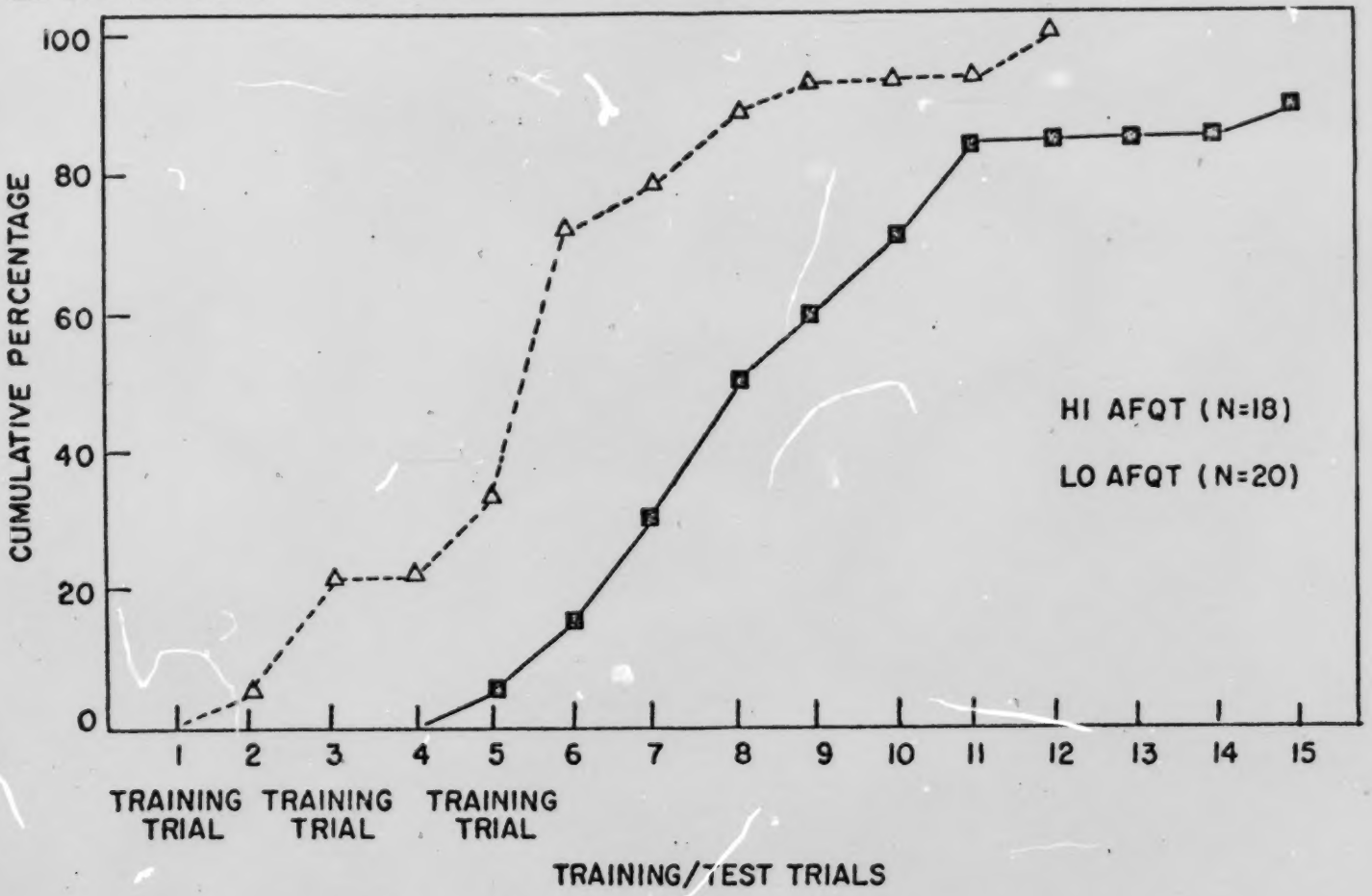



FIGURE 10

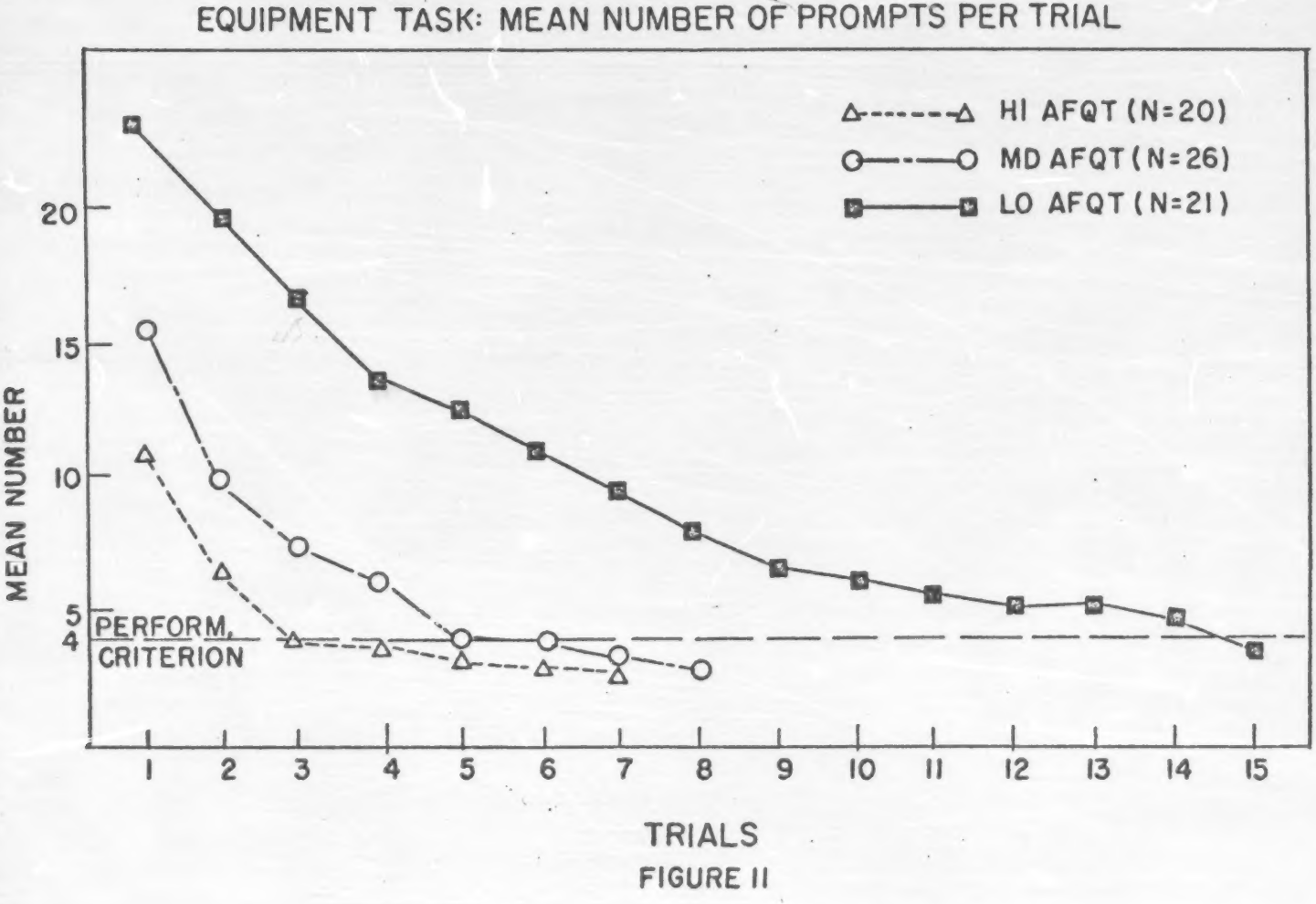

EQUIPMENT TASK: MEAN NUMBER OF PROMPTS PER TRIAL
HI AFQT (N=20)
MD AFQT (N=26)
LO AFQT (N=21)
MEAN NUMBER
20
15
10
5
4
PERFORM. CRITERION
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
TRIALS

FIGURE II

EQUIPMENT TASK: CUMULATIVE PERCENTAGE OF SUBJECTS REACHING CRITERION PER TRIAL

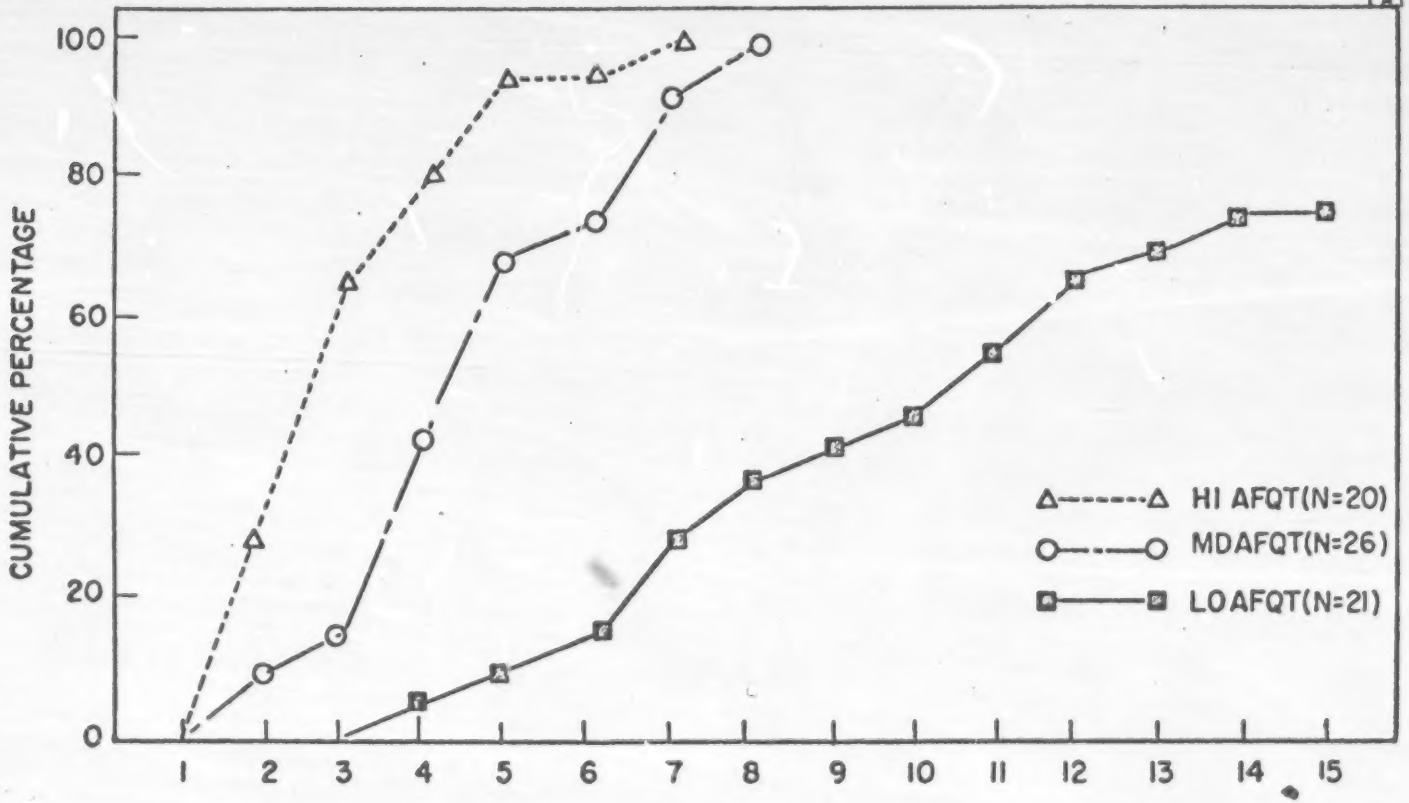



FIGURE 12

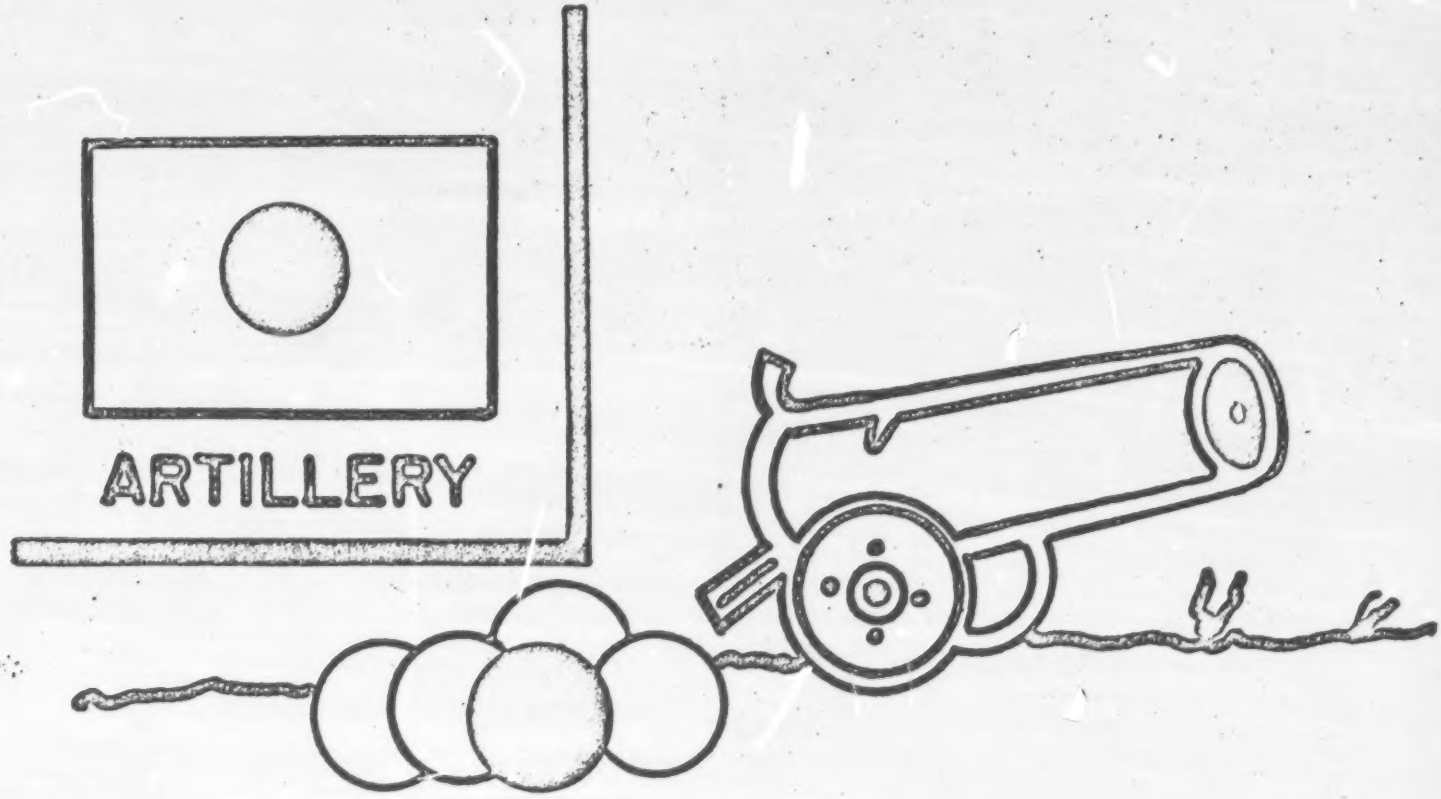



FIGURE 13

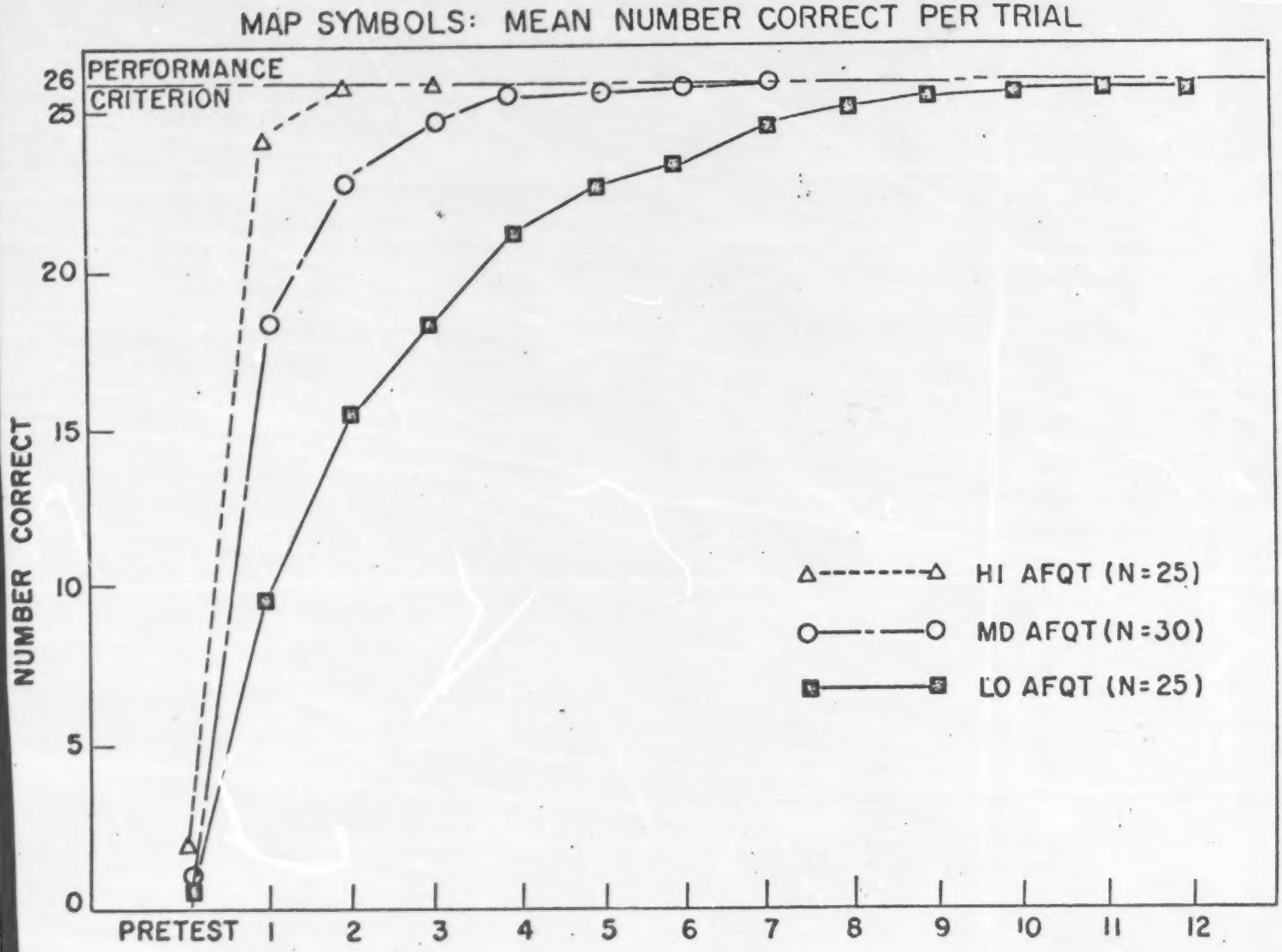

MAP SYMBOLS: MEAN NUMBER CORRECT PER TRIAL
PERFORMANCE CRITERION
26
25
20
15
10
5
0
NUMBER CORRECT
HI AFQT (N=25)
MD AFQT (N=30)
LO AFQT (N=25)
PRETEST
1
2
3
4
5
6
7
8
9
10
11
12
TRIALS

FIGURE 14

MAP SYMBOLS: CUMULATIVE PERCENTAGE OF SUBJECTS REACHING CRITERION PER T

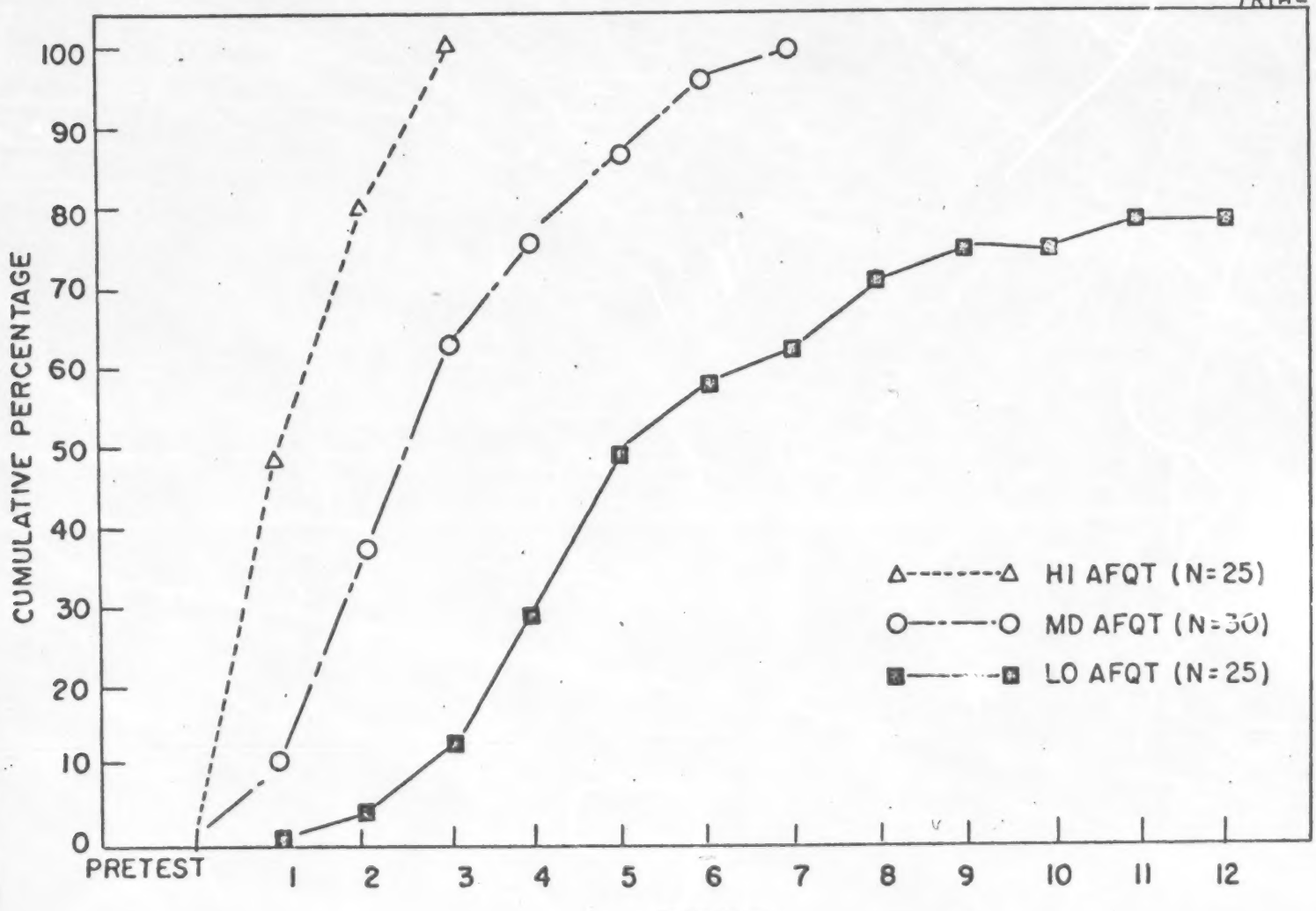



FIGURE 15

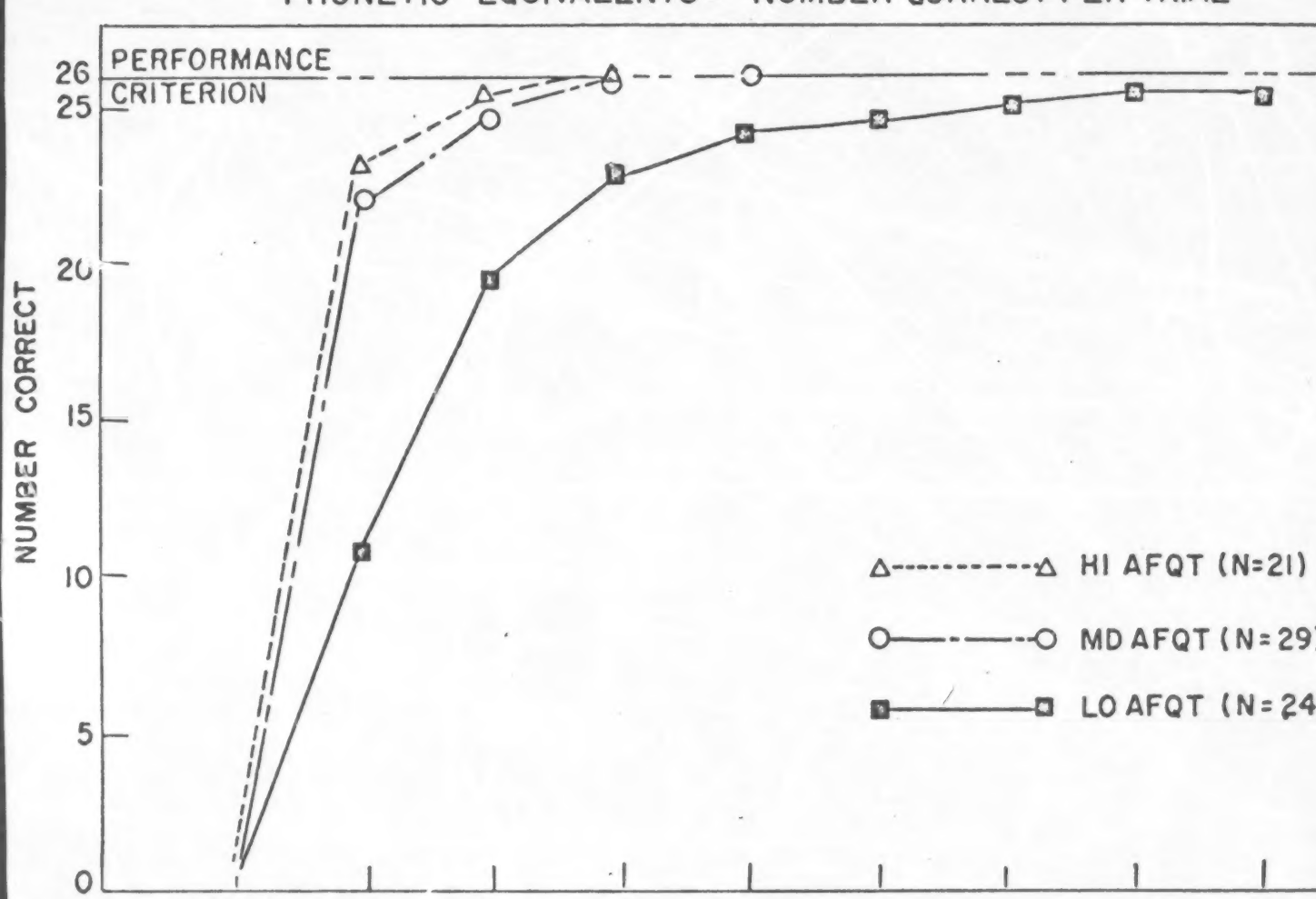

PHONETIC EQUIVALENTS: NUMBER CORRECT PER TRIAL
PERFORMANCE CRITERION
26
25
20
15
10
5
0
NUMBER CORRECT
HI AFQT (N=21)
MD AFQT (N=29)
LO AFQT (N=24)
PRETEST
1
2
3
4
5
6
7
8
TRIALS

FIGURE 16

PHONETIC EQUIVALENTS: CUMULATIVE PERCENTAGE OF SUBJECTS REACHING CRITERION PER TRIAL

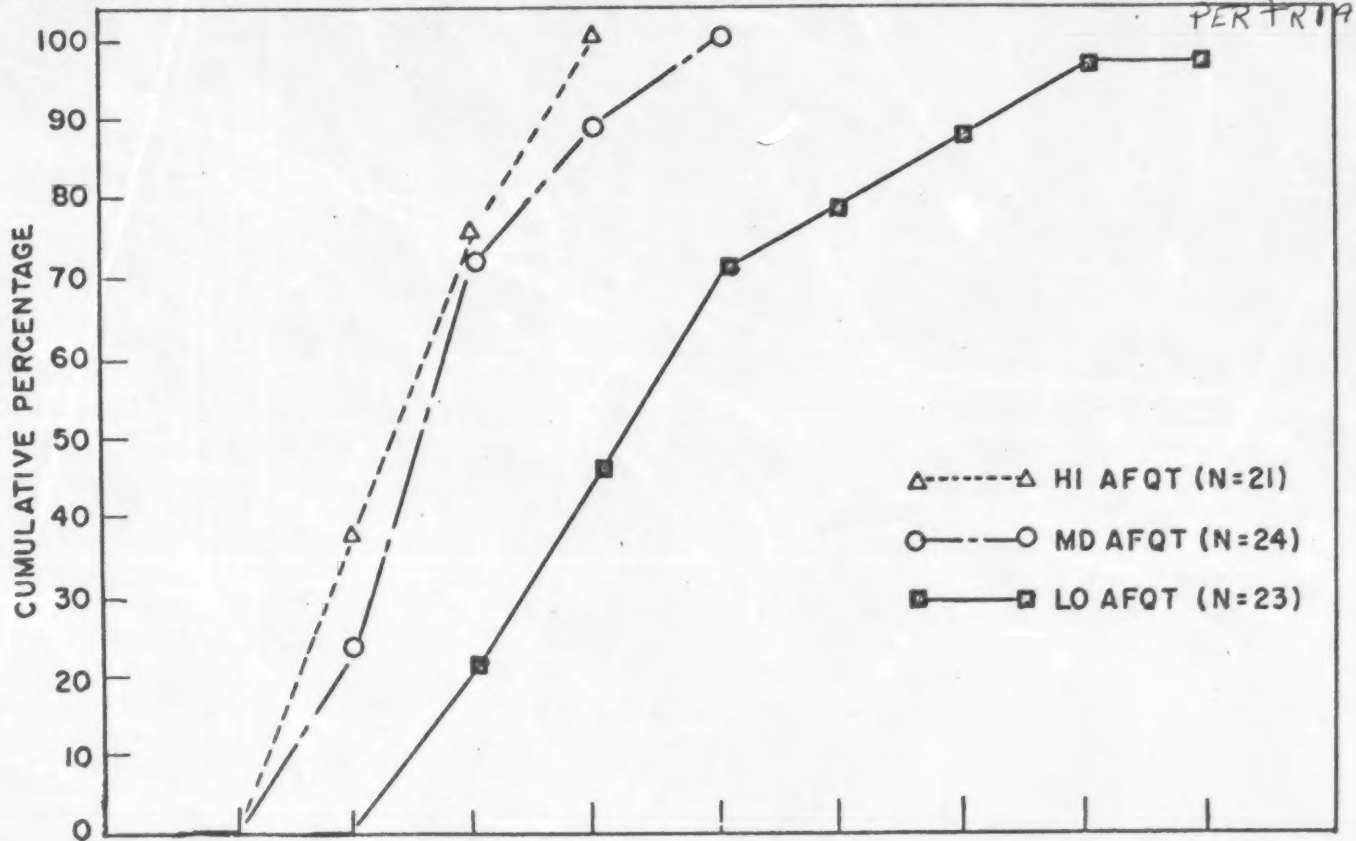
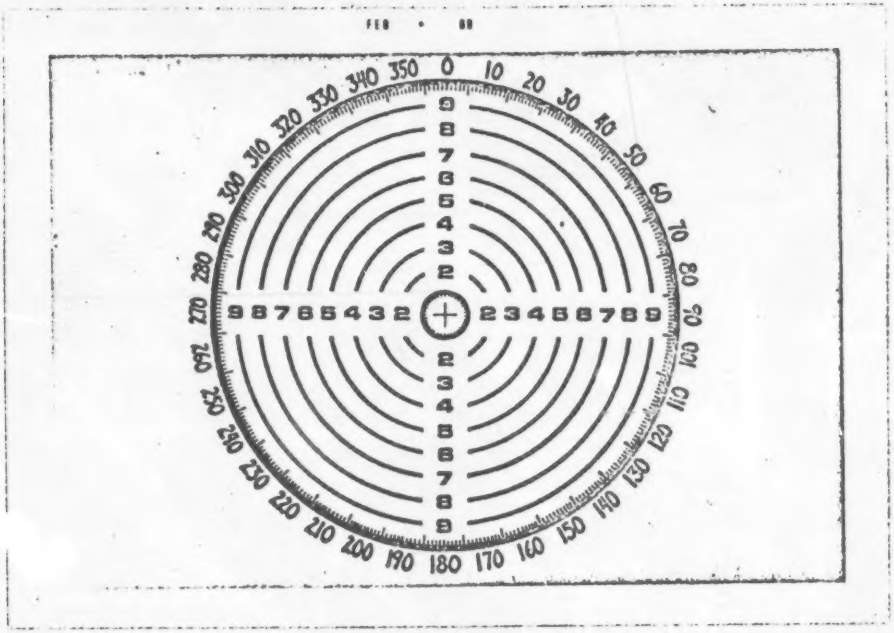



FIGURE 17




POSITION PLOTTING BOARD

FIGURE 18

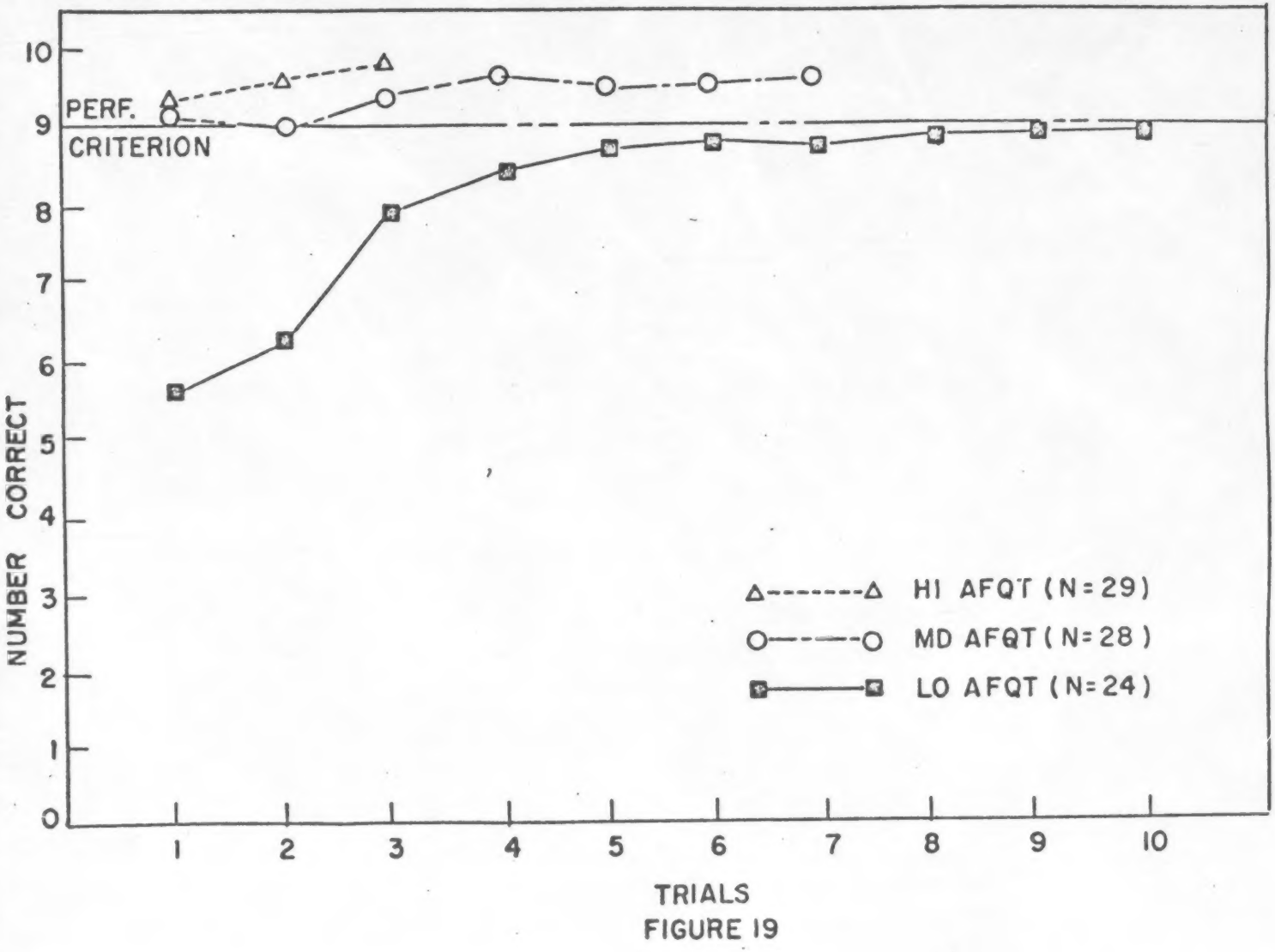

POSITION PLOTTING: MEAN NUMBER OF CORRECT PLOTS PER TRIAL
PERF. CRITERION
NUMBER CORRECT
TRIALS
HI AFQT (N=29)
MD AFQT (N=28)
LO AFQT (N=24)

FIGURE 19

POSITION PLOTTING: CUMULATIVE PERCENTAGE OF SUBJECTS REACHING CRITERION PER TRIA

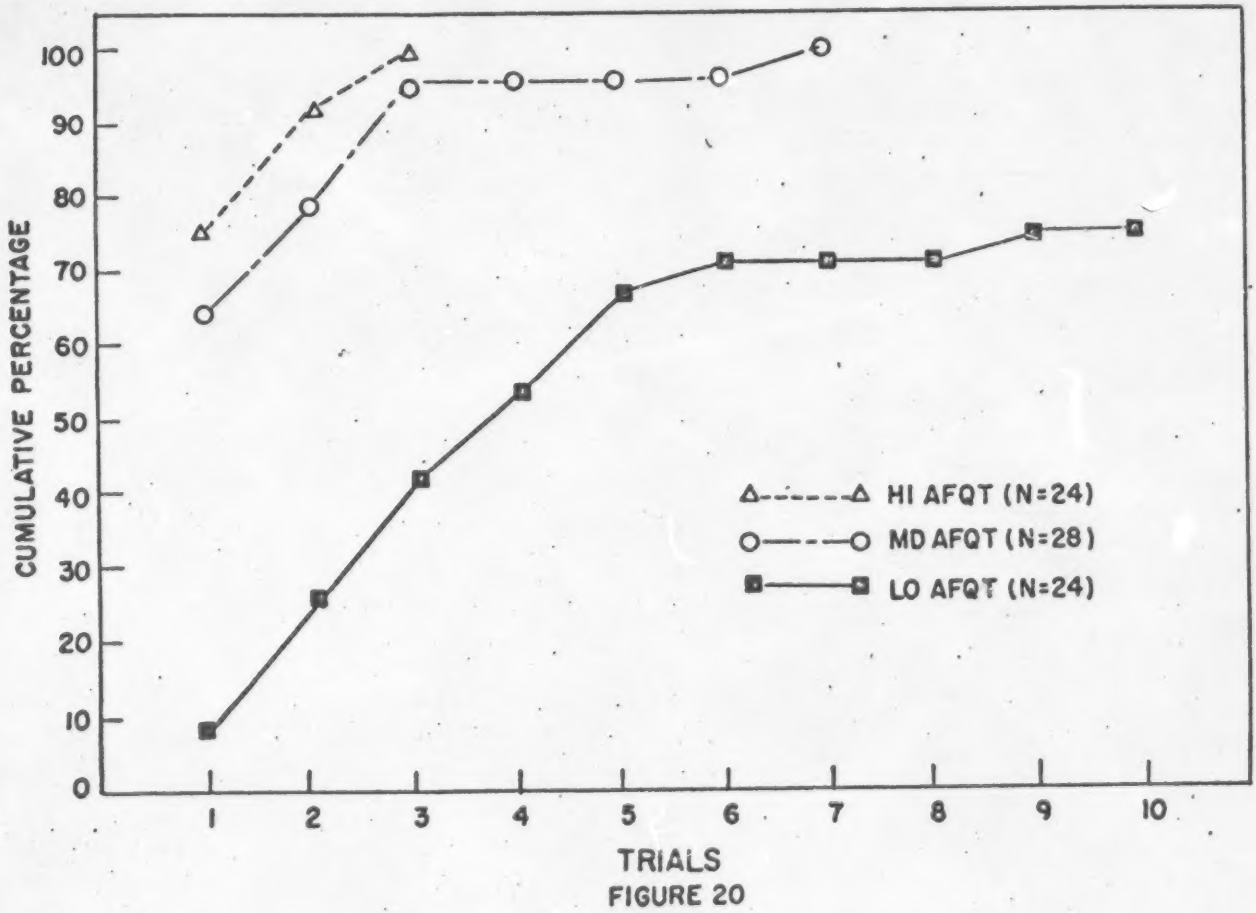



FIGURE 20

KEY WORDS

Acquisition of Military Knowledge and Skill

Aptitude

Armed Forces Qualification Test

Army Aptitude Areas

Army Classification Battery

Army Training Test 21-2

Basic Combat Training

Individual Differences

Marginal Manpower

Project 100,000

Rate of Learning

Task Complexity

Training Methodology

Training Performance

ABSTRACT

To assess the effects of wide differences in aptitude on the acquisition of knowledge and skills, a sample of 183 trainees was divided into three maximally distant aptitude groups on the basis of their AFQT scores: High aptitude, AFQT 90-99; Middle aptitude, AFQT 45-55; Low aptitude, AFQT 10-21. Each subject was individually trained to a performance criterion in differing combinations of a battery of eight tasks representative of vocational training. A variety of supplementary psychometric and scholastic achievement data was analyzed. The results were consistent in demonstrating large differences related to aptitude. As groups, high aptitude individuals excelled, low aptitude individuals did poorly, and middle aptitude groups fell in an intermediate range on all measures.

696 E5164

MAR 1969

Prefatory Note

This paper reports research conducted on Work Unit SPECTRUM, which is concerned with the development of efficient training across all aptitude levels. The research reported here was conducted at HumRRO Division No. 3, Recruit Training, Fort Ord, California.

CLEARED
FOR OPEN PUBLICATION
11 APR 1969 6
DIRECTORATE FOR SECURITY REVIEW (OASD-PA)
DEPARTMENT OF DEFENSE

001962